AL/OE-TR-1996-0178

UNITED STATES AIR FORCE
ARMSTRONG LABORATORY

PROCEEDINGS OF THE 1994 CONFERENCE ON TOXICOLOGY: CONFERENCE ON TEMPORAL ASPECTS IN RISK ASSESSMENT FOR NONCANCER ENDPOINTS

W. Bruce Peirano

ENVIRONMENTAL CRITERIA AND ASSESSMENT OFFICE
U. S. ENVIRONMENTAL PROTECTION AGENCY

David R. Mattie

OCCUPATIONAL AND ENVIRONMENTAL
HEALTH DIRECTORATE TOXICOLOGY DIVISION
ARMSTRONG LABORATORY
WRIGHT-PATTERSON AFB OH 45433-7400

Patricia M. Fleenor

MANTECH GEO-CENTERS JOINT VENTURE
P. O. BOX 31009
DAYTON, OH 45437

March 1996

DTIC QUALITY INSPECTED 2

Occupational and Environmental Health Directorate
Toxicology Division
2856 G Street
Wright-Patterson AFB OH 45433-7400

Approved for public release; distribution is unlimited.

19990427 009

NOTICES

When US Government drawings, specifications or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Please do not request copies of this report from the Air Force Armstrong Laboratory. Additional copies may be purchased from:

National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22161

Federal Government agencies and their contractors registered with the Defense Technical Information Center should direct requests for copies of this report to:

Defense Technical Information Service
8725 John J. Kingman Rd., Ste 0944
Ft. Belvoir, Virginia 22060-6218

DISCLAIMER

This Technical Report is published as received and has not
been edited by the Technical Editing Staff of the Air Force Armstrong Laboratory.

TECHNICAL REVIEW AND APPROVAL

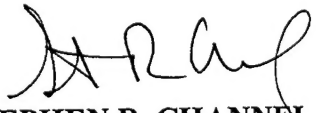
AL/OE-TR-1996-0178

The animal use described in this study was conducted in accordance with the principles stated in the "Guide for the Care and Use of Laboratory Animals", National Research Council, 1996, and the Animal Welfare Act of 1966, as amended.

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE DIRECTOR



STEPHEN R. CHANNEL, Maj, USAF, BSC
Branch Chief, Operational Toxicology Branch
Air Force Armstrong Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 1996	Interim Report - 18-20 April 1994

4. TITLE AND SUBTITLE
Proceedings of the 1994 Conference on Toxicology: Conference on Temporal Aspects in Risk Assessment for Noncancer Endpoints

5. FUNDING NUMBERS
Contract F41624-96-C-9010
PE 62202F
PR 7757
TA 7757A0
WU 7757A001

6. AUTHOR(S)
W.B. Peirano, D.R. Mattie, and P.M. Fleenor

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
ManTech Geo-Centers Joint Venture
P.O. Box 31009
Dayton, OH 45437

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Armstrong Laboratory, Occupational and Environmental Health Directorate
Toxicology Division, Human Systems Center
Air Force Materiel Command
Wright-Patterson AFB, OH 45433-7400

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AL/OE-TR-1996-0178

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The Conference on Temporal Aspects in Risk Assessment for Noncancer Endpoints was a continuation on a theme of risk assessment and toxicology that began in April, 1993, with the Conference on Risk Assessment Paradigm After Ten Years: Policy and Practice Then, Now, and in the Future. The purpose was to provide opportunities for research scientists, risk assessment practitioners, end users of risk analyses, and those interested in learning more about the rapidly evolving field of risk assessment to interact and to evaluate the state of the art of risk assessment. The goals were (1) recognition of temporal issues in the risk assessment process, (2) examination of critical time factors in hazard identification, dose-response, and exposure characterization, and (3) exploration of limitations in risk characterization.

The conference was cosponsored by the Tri-Service Toxicology Consortium (Air Force, Army, and Navy), Environmental Protection Agency, and Agency for Toxic Substances and Disease, and was attended by representatives of government, industry, and academia.

14. SUBJECT TERMS
Temporal factors; risk assessment; dose-response; endpoints; cellular toxicity

15. NUMBER OF PAGES
254

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Standard Form 298 (Rev. 2-89) (EG)
Prescribed by ANSI Std. 239.18
Designed using Perform Pro, WHS/DIOR, Oct 94

THIS PAGE INTENTIONALLY LEFT BLANK.

PREFACE

Attended by over 200 representatives of government, industry, and academia in the fields of toxicology and risk assessment, the conference "Temporal Aspects in Risk Assessment for Noncancer Endpoints" was featured sessions on

- Temporal Factors of Exposure in Identifying Hazards
- Effects of Exposure Patterns on Dose/Response
- Exposure Assessment and the Life Cycle Timeline
- Characterization of Risk in a Temporal Context

The papers in this volume span a wide range of topics that should be of interest to individuals in the fields of toxicology and risk assessment.

We would like to thank Lois Doncaster, conference coordinator, as well as the support staff of ManTech Environmental Toxic Hazards Research Unit for the preparation and coordination of the conference; the technical editors for reviewing the manuscripts, and sponsoring agencies for their assistance.

The full proceedings of this conference has been published by the journal *Inhalation Toxicology* in the August 1995 issue (Vol. 7: #6).

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK.

INTRODUCTION

W. Bruce Peirano
Environmental Criteria and Assessment Office
U.S. Environmental Protection Agency

David R. Mattie
Toxicology Division
Armstrong Laboratory
U.S. Air Force

Patricia M. Smith
Toxic Hazard Research Unit
ManTech Environmental Technology, Inc.

The Conference on Temporal Aspects in Risk Assessment for Noncancer Endpoints" was held at the Hope Hotel and Conference Center located at Wright-Patterson Air Force Base, Ohio, April 18–20, 1994. This conference was a continuation on a theme of risk assessment and toxicology that began in April of 1993 with the Conference on "The Risk Assessment Paradigm After Ten Years: Policy and Practice Then, Now, and in the Future". This purpose of this conference, like its predecessor conferences, was to provide opportunities for research scientists, risk assessment practitioners, end users of risk analyses, and those interested in learning more about the rapidly evolving field of risk assessment to interact and to evaluate the state-of-the-art of risk assessment. This conference's uniqueness is due, in part, to the cosponsorship by the Tri-Service Toxicology (U.S. Air Force, Army and Navy), Wright-Patterson Air Force Base; the Office of Research and Development, U.S. Environmental Protection Agency; and the Division of Toxicology, Agency for Toxic Substances and Disease Registry; with the cooperation of the National Research Council Committee on Toxicology.

The 1994 conference theme of temporal aspects in risk assessment of noncancer endpoints was a result of an identified need for a central forum where temporal and noncancer-related aspects could be presented and discussed. This evolved from a growing realization of the importance and impacts that temporal issues play in assessing occupational/environmental risks.

The goals for the conference were:

- Recognition of temporal issues in the risk assessment process
- Examination of critical time factors in hazard identification, dose-response, and exposure characterization

- Exploration of limitations in risk characterization

The introductory Session provided an overview of the conference, highlighted in this journal issue by the examples presented in the paper by Donald Mattison. Session II, "Temporal Factors of Exposure in Identifying Hazards," covered aspects of critical periods of exposure for inducing reproductive, developmental, immunological, neurological, and pulmonary hazards. Session III, "Effects of Exposure Patterns on Dose/Response," covered topics such as, what is meant by the term "dose;" the effect of multiple route exposure patterns; use of biomarkers to determine dose; and assessing exposure patterns in dose extrapolation across time. Session IV, "Exposure Assessment and the Life Cycle Timeline," covered topics on assessing exposure from bioaerosols during routine operations; exposure assessment of male reproductive effects; exposure characterization of the Kuwait oil well fires; exposure assessment at National Priority List sites; and refinements in the exposure assessment process. The final session, "Characterization of Risk in a Temporal Context," presented examples on assessing health risks to occupationally exposed Navy personnel; time-response profiles on pulmonary effects; temporal aspects of pesticide exposure; and temporal aspects of pediatric lead toxicity.

The selected conference papers published in this issue collectively reflect the growth in our understanding that there is more to risk assessment than chronic cancer assessments and time-weight averaging the exposures for dose–response modeling. The subtleties, such as time critical exposures in assessing developmental effects and preexposures in immunological effects, are increasingly being realized in assessments of risk and becoming the bases for risk based decisions. The platform and poster conference presentations have already generated much interactive discussion that will undoubtedly result in new research and methods related to temporal risk assessment issues. The goal of publishing the papers found in this issue is not only to disseminate information but also to further facilitate thoughts, discussions and new directions in this multidisciplinary field. Further synergetic activity is anticipated during next year's conference on "Risk Assessment Issues for Sensitive Populations."

We would like to thank the authors for contributing a written document as well as presenting at the conference. We would like to thank the conference session co-chairs David Macys, Kathryn Mahaffey, Daniel Caldwell, Bob Sonawane, Jewell Wilson, John Wyman, Warren Jederberg, Michael Callahan, Richard Thomas, Dave Reisman, John Latendresse, and Rogene Henderson for the time and expertise they expended in organizing and carrying out this conference. We would also like to thank the many

peer reviewers of these papers for their many hours of diligent work; and Shelia Brooks and Betsy Huber of ManTech Environmental for their invaluable assistance in coordinating the review, compiling, and editing processes that resulted in this proceedings. Finally, thanks to the conference attendees and readers of this journal issue for their involvement and willingness to help push the frontiers of risk assessment.

SESSION I

INTRODUCTORY SESSION

TEMPORAL ISSUES IN REPRODUCTIVE RISK ASSESSMENT[1]

Sati Mazumdar and Donald R. Mattison
Graduate School of Public Health, University of Pittsburgh

C.V. Damaraju
Graduate School of Public Health, University of Pittsburgh and Schering-Plough Research Institute, Kennilworth, NJ

ABSTRACT

While there is increasing interest in characterizing reproductive hazards, reproductive endpoints are complex, difficult to characterize and variable across populations and across time in the same individual. It is generally appreciated that time is a critical variable, however, it is seldom fully considered in the characterization of reproductive risks. For example, female fecundity varies with cycle day, and while both male and female fecundity are dependent on historical performance they also vary over time in a given individual. In addition, sensitivity of the male and female reproductive systems to toxicants varies during ovarian and spermatogenic cycles. Initial characterization of toxic effects on fecundity have considered the time necessary to achieve a particular endpoint, such as time to pregnancy. However, analysis of time in the calculation of benchmark doses has not been considered for quantitative characterization of reproductive risks. This paper suggests an approach for characterizing risks to reproduction in the context of fecundity and fertility, reviews temporal factors in reproduction, and discusses an approach for the analysis of time dependent factors in the calculation of benchmark doses for reproductive toxicity. Total sperm counts from a rabbit inhalation experiment with dibromochloropropane were utilized to explore this approach. Three variance stabilizing transformations (cube root, logarithmic and an optimal transformation obtained using an additivity and variance stabilizing algorithm) were explored. Two longitudinal data analysis techniques using random effects models and generalized estimating equations were used to characterize the dose–response relationships and calculate specific benchmark doses for alteration of spermatogenesis by dibromochloropropane. These analyses suggest approaches for defining the most appropriate transformations to stabilize variance and methods for dealing with time in experiments designed to characterize reproductive response to toxicants. Similar analyses are necessary for other biomarkers of

1 This work was supported in part by a grant from the American Industrial Health/Hygiene Council (AIHC). We thank Drs. George Daston, Betsy Carlton, Willem Faber, Ronald Fliller, William Pease and Kun Jin for their valuable suggestions. Presented at the 1994 Conference on Temporal Aspects in Risk Assessment for Noncancer Endpoints, April 18-20, Dayton, OH.

male and female reproductive function before benchmark doses are calculated to protect human reproductive health.

Key words and Phrases: Dibromochloropropane (DBCP); sperm number; variance stabilizing transformations; longitudinal data analysis; reproductive risks; benchmark dose.

INTRODUCTION

Over the past decade, several approaches have been explored to characterize reproductive risks, including: (i) statistical approaches for dose response modeling, such as the no observed adverse effect level (NOAEL)-safety factor (SF) approach to determine the allowable daily intake (ADI), the benchmark dose (BD)-SF approach to determine the reference dose for reproductive toxicity (R_fD_R); and (ii) biologically based approaches including classical pharmacokinetics (PK), physiologically based pharmacokinetics (PBPK), and biologically based dose response modeling (BBDR). Analysis of these methods suggests that at the present time, statistical approaches which characterize the dose–response relationship are most conservatively applied to risk assessment for reproductive toxicity. However, as concern for the hazard increases, the development of biologically based methods will occur.

Reproductive toxicity studies conducted in experimental animals have an important bearing on risk assessment designed to protect human reproductive health (Mattison, 1991; Meistrich, 1992; Mattison et al., 1989, 1990). Acute, subchronic and chronic exposures to reproductive toxicants may impair reproductive function leading to subfertility or infertility (Scialli and Zinaman, 1993; Paul, 1993). Among the biological markers available to characterize male fecundity, measures of sperm counts (e.g., number of sperm in the ejaculate) are accessible, frequently studied, and potentially useful biomarkers of male reproductive toxicity for exploration of statistical techniques.

Studies in reproductive toxicity provide data, which can be used to investigate the underlying structure of the dose-response relationship between a biomarker such as sperm number and exposure to a given chemical. Animal studies have demonstrated this, and data from such studies have been used to identify and evaluate the mathematical forms of dose-response relationships (Krewski and Franklin, 1991). For subchronic and chronic studies, investigations based on a cross-sectional design present limited information for the exploration of a cumulative dose-effect on the response. Methods capable of assessing toxic effects on spermatogenesis, for example, require longitudinal study designs. Such experiments include a pre-exposure period, an exposure period which covers the duration of spermatogenesis (9–14 weeks), and a post-exposure or recovery period of similar duration. These

experiments have proven useful from a statistical standpoint in the understanding of toxicological phenomena that lead to impairment of fertility (Saegusa, 1989; Williams et al., 1990).

Risk assessment from findings in laboratory animals requires extrapolation to determine levels of allowable exposure for humans (Working and Mattison, 1993). One method proposed for extrapolating the dose-response relationships determined from animal studies to humans is via an interspecies extrapolation factor (Meistrich, 1989a, 1992). Another approach, as described above, calculates a BD, which is usually defined as the lower 95% confidence bound on the dose associated with a 10% change in effect from the control group, abbreviated here as $L_{95}ED_{10}$ (Crump, 1984; Pease et al., 1991). The BD can then be used to calculate the R_fD_n, the exposure at which little or no adverse reproductive toxicity is expected in the human population. The reference dose methodology is discussed in general by Barnes and Dourson, (1988).

The purpose of this paper is to present a couple based framework for reproductive risk assessment and describe several statistical approaches for the determination of time dependent dose-response relationships for selected reproductive endpoints. These approaches incorporate exploratory data analysis techniques to reduce variability in reproductive measures that are often encountered in such experiments and generally regarded as a hindrance to dose-response modeling. The data used in this analysis are from the total sperm count per ejaculate for rabbits taken from the fertility study, *"1,2-Dibromo-3-Chloropropane: Inhalation Fertility Study in Rats and Rabbits"* that was conducted by Rao and his colleagues (1980, 1982).

RISK ASSESSMENT FOR REPRODUCTIVE TOXICITY

Qualitative or quantitative risk assessment provides a formal methodology to protect the public from adverse health effects resulting from exposure to chemicals, physical factors or biological agents. While methods for characterizing risks to human health have evolved over the past four decades, the development of most quantitative risk assessment methods has occurred over the past ten to fifteen years (National Research Council, 1989), and has focused predominately on cancer. In addition to Risk Assessment, there are three other risk sciences: Risk Perception, Risk Communication and Risk Management, the discussion of which is beyond the scope of this paper. Risk assessment utilizes both scientific and quasi-scientific considerations and incorporates both qualitative and quantitative data. As a result, risk assessments are often a matter of contention (Jasanoff, 1990).

Risk assessment involves four steps; hazard identification, hazard characterization, exposure assessment and risk characterization. The nature of each step and the approach employed to combine the data into the final risk assessment must be based on our best understanding of the biology of the process for which the risk assessment is being conducted (Mattison, 1991; Figure I-1, Table I-1). It should be noted that we are using an expanded definition here for risk assessment that allows for the inclusion of site and mechanism of action information.

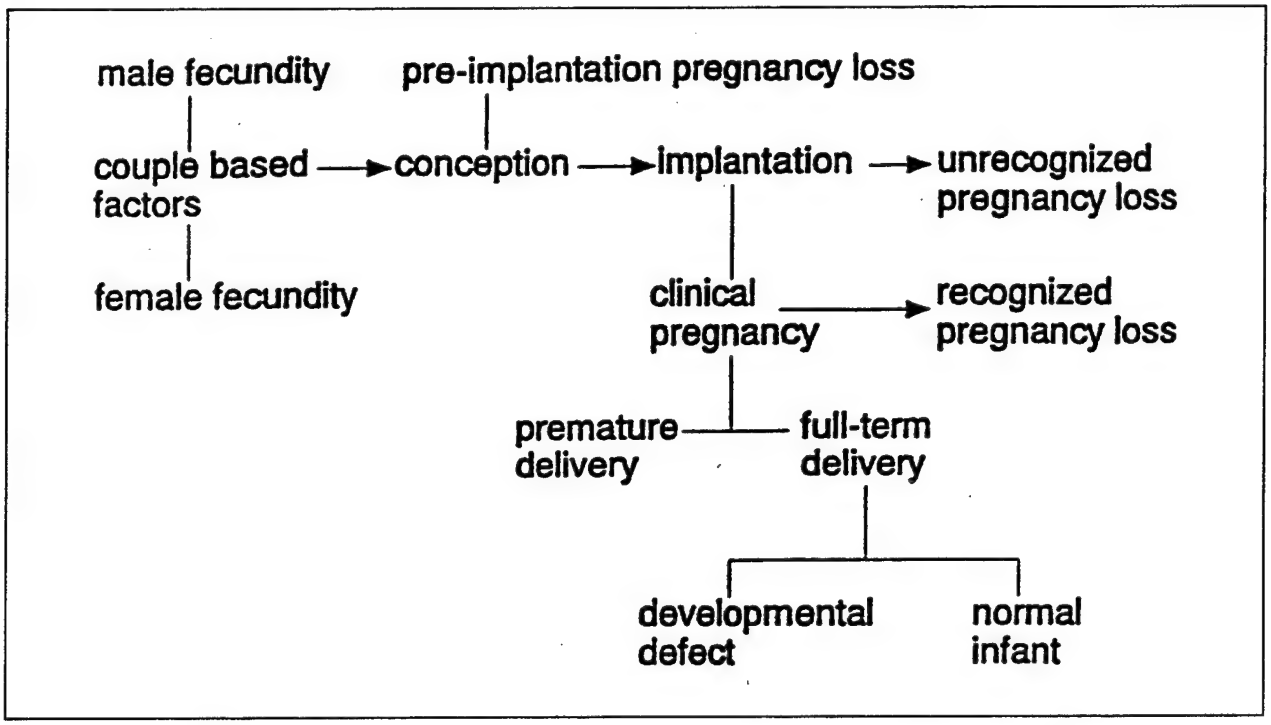


Figure I-1. Couple-Based Model for Biomarker-Based Risk Assessment of Reproductive Toxicity.

Table I-1.

I

ndices of Fecundity in Laboratory Animals[a].

Indices of Male Fecundity	Indices of Female Fecundity
Disruption of seminiferous epithelium	Estrous cycle disruption resulting in anovulation
Alterations in gonadal function causing decreased testicular spermatid number or decreased sperm count in the epididymis, vas deferens, or ejaculate	Significant reduction in the number of ovarian follicles/oocytes Altered uterine histology
Decrease in percentage of motile spermatozoa	Altered ovarian histology characterized by reduced corpora lutea or increased number of ovarian cysts
Significant change in sperm morphology	Altered concentration or temporal patterns of testosterone, luteinizing hormone (LH) or follicle-stimulating hormone (FSH)
Alterations in reproductive organ weight (e.g., testes, epididymides, seminal vesicles, or prostate)	Alterations in ovarian or uterine weight
Altered concentration or temporal patterns of testosterone, luteinizing hormone (LH) or follicle-stimulating hormone (FSH)	Delayed puberty
	Premature reproductive senescence

[a] Adapted from Mattison et al., 1990, *Reprod. Toxicol.* 4:163–165.

Statistical Approaches for Risk Assessment for Reproductive Toxicity

NOAEL-SF-ADI Approach: Initially the focus for risk assessment for reproductive toxicity converged on the identification of the NOAEL and the application of a SF (which ranged from 1 to 1000) to the NOAEL to determine the ADI (Barnes and Dourson, 1988; Crump, 1984; Mattison et al., 1989). This approach assumes thresholds for reproductive toxicity and presumes once the threshold is defined it is possible to set a safe level for human exposure. Although this approach is simple, it has definite disadvantages. To determine the NOAEL, some investigators have simply identified the highest dose for which the response is not statistically different from the control. Others have used an approach to identify the highest dose that does not have a statistically significant trend for a dose–response relationship. While these approaches are conceptually simple, they present some analytical problems to the risk assessor; the first method does not use all of the data, and in both methods the NOAEL is dependent on the number of animals in treatment and control groups. As the number gets smaller the NOAEL increases. These characteristics tend to reward poorly designed studies such as inappropriate selection and spacing of doses, small number of animals per treatment group, etc.

Benchmark Dose-SF-RfD Approach: This approach uses individual animal responses at all points in the dose–response curve to calculate a BD which is in or near the observed treatment range (e.g., a

lower confidence bound on ED_{10}). The BD is then divided by a SF (which can range from 1 to 1000) to determine the R_fD. While this approach has great appeal because it uses all of the data from a reproductive toxicity experiment, some have avoided this approach because of the statistical analysis required and the general uncertainty about how to apply a new and untested method.

Biologically Based Approaches for Risk Assessment

A range of assumptions are necessary to perform risk assessments for human health. These include extrapolations necessary from the high doses at which animal experiments are conducted to the low doses typical of human exposure. Other extrapolations include those from animals to humans and across route of exposure. As the risk sciences have developed improved methods for characterizing human health risks, the techniques have utilized more detailed biological information to decrease the uncertainty about these extrapolations (Aafjes et al., 1950; Amann and Howards, 1980; Amann and Hamerstedt, 1980; CECOS et al., 1982; David et al., 1979; Ewing and Mattison, 1987; Faustman et al., 1989; Foote et al., 1986ab; Guerrero and Rojas, 1975; Horning et al., 1981; Mattison, 1990, 1991, 1993; Mattison and Thomford, 1989; Mattison et al., 1990, 1991). A general consensus has developed which suggests that *risk assessments should be constructed in the context of the biological principles that govern the endpoint evaluated.* Figure I-1 summarizes some of these principles for reproductive toxicity. Factors which can influence reproductive toxicity include: male fecundity, dose and duration of exposure, stages of spermatogenesis effected by the chemical, distribution of the chemical to the testis and other male reproductive organs, mechanisms of action of the chemical, and the frequency and timing of intercourse with respect to treatment.

Comparison of Reproductive Risk Assessment with Traditional Risk Assessment

Risk assessment for reproductive toxicity is outside the general framework of risk assessment for cancer or end organ toxicity. A traditional risk assessment completes the four steps within the context of an exposure and effect characterized for an individual. However, in risk assessment for reproductive toxicity the male may be exposured and the response measured in the male, female or offspring. This complicates risk assessment for reproductive toxicity and suggests that the risk assessment process must be considered within a different (i.e., couple based) framework (Baird et al., 1986; Baird and Wilcox, 1985; Barrett, 1971; Barrett and Marshall, 1969; CECOS et al., 1982; David et al., 1979; Generoso et al., 1979ab; Generoso, 1980; Guerrero and Rojas, 1975; Levine et al., 1980, 1981; Mattison and Brewer, 1988; Mattison, 1991; Meistrich, 1984, 1988, 1989ab; Meistrich and Brown, 1983; Menken et al., 1986). In the context of the couple-based approach for characterizing the

reproductive risks, the endpoints which may be included are summarized on Figure I-1 and Table I-1. The risk assessment example included in this discussion will be conducted using a biological marker of male reproductive function — sperm number in the ejaculate.

BIOLOGICAL MARKERS FOR REPRODUCTIVE TOXICITY RISK ASSESSMENT

The utility of biological markers as tools for characterizing risk of reproductive toxicity following male exposure is not well defined. Biologic markers should be utilized to clarify relationships between exposure and reproductive disease (National Research Council, 1989). The availability of validated biological markers will allow the development of biomarker based risk assessment for reproductive toxicity. A biomarker based model of risk assessment should incorporate male fecundity (MF), female fecundity (FF) and couple fecundity (CF), all which may influence reproductive success.

Biomarkers of Male Fecundity

Biomarkers of male reproductive function characterize the ability of the male to fertilize the female and produce a normal offspring. Biomarkers which *may* be applicable for characterizing the impact of an exposure on male reproductive success include: anatomic factors (A_{fm}), ejaculate volume (E_v), ejaculate composition (E_c), sperm number per ejaculate (S_n), sperm motility (S_m), sperm morphology (S_s), and measures of the genetic integrity of the sperm (S_{gi}). These biomarkers contribute to a function which describes the role of male fecundity in reproductive success:

$$\text{Male Fecundity} = MF(Af_m, E_v, E_c, S_n, S_m, S_s, S_{gi}) \quad (1)$$

Selected biomarkers will be included in risk assessment for reproductive toxicity, depending on: availability of data, utility of a given biomarker, and the sensitivity of a reproductive endpoint to changes in a specific biomarker. Data from the National Toxicology Program (NTP) suggests that sperm concentration, motility and morphology may be reasonable biomarkers for characterizing male fecundity (Table I-2 and Morrissey 1989). The data in this study is limited, however, by the number of chemicals examined. There were 24 chemicals studied in a testing protocol designed to assess reproductive performance by continuous breeding (Lamb, 1989).

Table I-2. Utility of Selected Biomarkers as Predictors of Male Fecundity[a].

Biological Marker	Sensitivity (%)	Specificity (%)	Positive Pred. Value (%)	Negative Pred. Value (%)
Epididymal Weight	80	87	80	87
Testis Weight	62	83	80	67
Sperm Motility	69	92	90	73
Sperm Concentration	70	80	70	80
Abnormal Morphology (%)	60	79	67	73
Body Weight	38	58	50	47

[a] Data from Morrissey, 1989. *Toxicology of the Male and Female Reproductive Systems*, Hemisphere Publishing Corporation, New York. 199–216.

Biomarkers of Female Fecundity

These biomarkers characterize the ability of a female to be fertilized by a male and produce normal offspring. The biomarkers which *may* be applicable for characterizing female reproductive success include: anatomic factors (A_{ff}), ovulatory frequency (O_f), follicular phase characteristics (F_f), luteal phase characteristics (L_f), endometrial function (Ef), tubal function (T_f), DNA repair characteristics of the oocyte (O_{dna}) and genetic characteristics of the oocyte (O_{gi}). As described for the male, these biomarkers contribute to a function which describes the role of female factors in reproductive success:

$$\text{Female Fecundity} = FF(A_{ff}, O_f, F_f, L_f, E_f, T_f, O_{dna}, O_{gi}) \quad (2)$$

Use of an individual biomarker is dependent upon: availability of data and the utility of that biomarker for the reproductive endpoints under consideration.

Biomarkers of Couple Fecundity

These biomarkers characterize the ability of a given couple (or breeding pair for experimental animals) to produce offspring. The biomarkers which *may* be applicable for characterizing the role of couple specific factors in reproductive success include: frequency of intercourse (F_i), male-female interactions (M_f), and female-male interactions (F_m). These biomarkers contribute to a function which describes the role of couple specific factors in reproductive success:

$$\text{Couple Fecundity} = CF(F_i, M_f, F_m) \quad (3)$$

As described for MF and FF, biomarkers which characterize couple function will be included in risk assessment depending on: availability of data and utility of a given biomarker to predict the endpoint of concern.

Reproductive Risk Assessment

In the context of the couple-specific approach, reproductive risk is a function of individual and couple-specific factors and is represented by the equation:

$$\text{Reproductive risk} = RR\ (MF, FF, CF) \tag{4}$$

Note that the function describing reproductive risk is likely to be specific to the endpoint considered. For example, the function describing the male contribution to preimplantation pregnancy loss is likely to be different from the function describing the male contribution to infertility.

Animal data is needed to investigate relationships between biomarkers suggested to quantitate the impact of exposures to the male and the risk for reproductive toxicity. Methods to extrapolate this data from animals to humans, from high to low dose and across other factors are also required. This process will be illustrated with data from an experiment in which animal exposure to DBCP and alterations in sperm number in the ejaculate were quantitated. Analysis of NTP data from reproductive effects by continuous breeding (Lamb, 1989) suggests that morphologically normal sperm number in the ejaculate is a good predictor of male fecundity (Table I-2 and Morrissey, 1989). Using data from an experiment with DBCP it is possible to explore the effect on reproductive function quantitatively (Rao et al., 1980, 1982).

ASSESSING REPRODUCTIVE RISKS WITH LONGITUDINAL DATA ANALYSIS TECHNIQUES

Occupational exposure to pesticides is a risk factor for reproductive disease. Effects of reproductive toxicants like the soil fumigant *dibromochloropropane* (DBCP) (Whorton et al., 1977, 1979; Milby and Whorton, 1980; Rao et al., 1982; Potashnik, 1983; Goldsmith et al., 1984; Potashnik and Abeliovich, 1985; Eaton et al., 1986; Pease et al., 1991), *ethylene dibromide* (EDB, Dobbins, 1987; Ratcliffe et al., 1987; Williams et al., 1991), *Agent Orange* (Paul, 1993), and *alkylating agents* on humans or other species have been reported in the reproductive toxicology literature. These have documented dose and time dependent changes in semen characteristics such as sperm number, motility and morphology (Gladen et al., 1991), and damage to testicular function (Rao et al., 1980).

Statistical approaches for analysis of data from these toxicological studies is generally confined to establishing dose-response relationships using cross-sectional study designs. However, Williams et al. (1990) demonstrated that the inclusion of a pre-exposure period in the study design and utilization both between and within individual variations, increased the statistical power considerably. Moreover, to estimate the change in some response parameter that may be time-dependent, longitudinal data analysis techniques are considered more appropriate than cross-sectional data analysis (Ware et al., 1990). Analytical techniques utilizing the longitudinal nature of the response measures should therefore be employed to assess the reproductive toxicity of DBCP.

DBCP Inhalation Rabbit Fertility Study

Groups of 10 male New Zealand white rabbits were exposed to DBCP vapor in individual animal chambers at four different dose levels: 0, 0.1, 1.0, and 10 ppm respectively. The pre-exposure data collection period was two weeks. DBCP exposure lasted for 14 weeks with a 5 days/week-6 hours/day schedule with the exception of rabbits in the 10 ppm group that were exposed for only 8 weeks due to a high incidence of mortality (Table I-3). The post exposure period began after week 14 and extended through week 46. The duration of exposure and recovery periods were designed to encompass the length of the spermatogenic cycle in the rabbit. Semen specimens were collected from the rabbits prior to exposure, on a weekly basis during the 14 week exposure period, and at periodic intervals during recovery. Endpoints of semen evaluation consisted of volume, sperm concentration, sperm motility, and viability.

Rao et al., (1980) reported treatment-related alterations in male reproductive function and deleterious effects on the reproductive capacity of rabbits treated with 1 or 10 ppm of DBCP for 14 weeks. No adverse effects were observed in rabbits exposed to 0.1 ppm following conventional methods of evaluation. However, electron microscopy did show equivocal effects in the rabbits exposed to 0.1 ppm immediately following the 14 week exposure period, but not after the recovery period. It is significant to note that male rabbits displaying moderate testicular atrophy and other DBCP induced alterations were still fertile and, upon breeding with unexposed females, were able to sire normal appearing pups. It was concluded that the NOAEL for DBCP in inhalation treatment in rabbits was 0.1 ppm, except for an equivocal effect on sperm morphology detected only by electron microscopy, which did not appear to affect fertility.

Table I-3. Number of Rabbits in the DBCP Inhalation Study[a].

Week	ppm DBCP			
	0	0.1	1	10[b]
Pre-exposure				
−2	9	8	9	8
−1	9	8	6	6
Exposure				
1	10	10	9	9
2	10	7	10	7
3	9	10	9	10
4	10	10	10	10
5	10	10	10	10
6	10	10	9	9
7	9	10	9	7
8	9	9	10	6
9	7	6	7	6
10	8	10	10	5
11	8	10	7	4
12	7	9	7	4
13	9	10	9	5
14	8	9	7	5
Post-exposure				
16	6	6	5	5
19	6	5	4	5
24	6	6	5	5
26	5	4	5	5
27	6	6	5	5
28	5	6	5	4
30	5	6	5	5
32	6	4	5	5
34	6	6	4	5
36	6	6	5	5
38	6	6	5	5
40	5	6	4	5
42	5	6	4	5
44	5	6	3	5
46	5	6	3	5

[a] Data from Rao et al. (1980). All measurements from animal 34 in the 10 ppm group were discarded. Observations are missing due to death, sacrifice and insufficient sample or low sperm count.

[b] Because of mortality and morbidity the exposure period for rabbits in the 10 ppm group ends with week 8 and the post-exposure period begins in week 9.

This rabbit inhalation dataset has also been analyzed by four alternative approaches for deriving regulatory levels for exposure to reproductive toxicants (Pease et al., 1991). These approaches were: (i) Proposition 65 approach; (ii) the Environmental Protection Agency (EPA) approach; (iii) the Benchmark Dose (BD) approach; and (iv) the Quantitative Risk Estimation (QRE) approach (Meistrich, 1992). Pease et al., (1991), assumed a regression model with dose as the independent variable and the mean sperm concentration from weeks 11 through 14 (for all dose groups except the 10 ppm dose group) as the dependent variable, and obtained a BD estimate of 0.015 mg/kg-day which was used to calculate a R_fD_{rt}. This BD estimate is also referred to as $L_{95}ED_{10}$ and is defined as the 95% lower confidence limit on the dose level producing a 10% change in the response from background. The background total sperm count in the ejaculate was taken as 400 million.

It is important to note that these approaches have different data requirements. Neither the Proposition 65 nor the EPA approach provides information about the extent of health risk remaining at a given exposure level. For the Proposition 65 approach, a NOAEL is identified and divided by a 1000-fold SF to calculate a R_fD_{rt}. For the EPA approach a NOAEL is identified and divided by a selected SF (which may be as large as 1000-fold). In the QRE approach, uncertainties in intra- and inter-species extrapolation and dosing regimen are addressed during the development of a potency estimate for the toxicant. In contrast, the BD approach provides estimates of the magnitude of the sperm count reduction at the regulatory level.

Statistical Analyses

Several statistical approaches were considered in the present analysis.

Regression approach: The time dependence of the dose parameters was examined by fitting a linear regression model with indicator or dummy variables for each week. The model is given by

$$y_{ij} = \alpha+\beta d_j + \gamma_1 D_1 +..+ \gamma_{k-1} D_{k-1} +\beta_1 d_j D_1 +..+\beta_{k-1} d_j D_{k-1}+e_{ij}, \quad (5)$$

$i = 1, 2,...,N_j$ and $j = 1,2, 3, 4$. Here, y_{ij} is the sperm measure of the ith animal at the jth dose level, d_j, N_j is the number of animals at the jth dose level, α and β are the intercept and slope for the baseline week, D's represent (k–1) dummy variables for k weeks of data, $\gamma_1, \gamma_2,...,\gamma_{k-1}$ and $\beta_1, \beta_2,...,\beta_{k-1}$ are additional coefficients to include time and dose x time interactions and e_{ij} are errors (Myers, 1990).

Exploratory Data Analysis for a Transformation via Additivity and Variance Stabilization: This analysis employed a technique for nonparametric estimation of transformation of variables in a regression problem via additivity and variance stabilization (Tibshirani, 1988). The resultant algorithm

is called AVAS (Additivity and Variance Stabilization) (S-PLUS, 1992). This method uses the asymptotic variance-stabilizing transformation for the response variable. Given two random variables X and Y, the goal is to find real-valued, measurable transformations s(X) and g(Y), such that

$$E(g(Y)/ X = x) = s(x),$$

$$\text{var}(g(Y)/s(X)) = \text{constant.} \quad (6)$$

The transformation g(Y) is assumed to be strictly monotonic, and without loss of generality it is assumed to be strictly increasing. These transformations are sought as the postulated model is

$$g_0(Y) = s_0(X) + \varepsilon, \quad (7)$$

where $g_0(Y)$ is strictly increasing, and ε has a mean of 0 and is independent of X.

The response from control animals was considered to be Y and the time variable(week) as X for arriving at a parametric form for the AVAS transformation for Y. Constraining the transformation on time to be linear, AVAS transformed individual animal sperm measures over time were obtained. Using plots of these transformed values, a parametric form of the transformation was chosen by visual inspection and fitted. The resulting fitted model was used to determine the AVAS transformed values of the sperm measures in the other three dose groups.

Dose-Response Modeling with Longitudinal Data Analysis Techniques

In observational and laboratory experiments, missing data are quite common. For example, it is frequently not possible to account for all measurements as some animals die or sacrificed early because they are moribund. Two longitudinal data analysis techniques allowing correlations of multiple measures on the same animal and missing observations are used in the present study.

Restricted Maximum Likelihood (REML) Method for the Random-Effects Models

A convenient approach for the analysis of missing and/or unbalanced longitudinal data is the use of two-stage random effects models (Laird and Ware, 1982). The methodology consists of a two-stage modeling approach where the probability distributions for the response vectors of different individuals belong to a single family, but some parameters are allowed to vary across individuals (random effects) with a distribution specified at the second stage.

A dose–response model which is linear in both time and dose is given by

$$y_{ij} = a_0 + a_1 t_j + a_2 d_i + b_{0i} + b_{1i} t_j + b_{2i} d_i + e_{ij}, \quad (8)$$

where y_{ij} is an appropriate transform of the sperm measure of the ith individual at the jth time t_j with dose level d_i, a_0, a_1, and a_2 are the fixed effects parameters, b_{01}, b_{1i} and b_{2i} are random effects (individual deviations from the fixed effects parameters a_0, a_1 and a_2 respectively) and the e_{ij}'s are error terms. In this formulation i = 1, 2,..., N, where, N is the total number of animals, j represents the time when a particular measurement is taken and d_i is the ith dose level.

Therefore, the mean (population) model at time t_j and dose level d_i is given by

$$E[y_{ij}] = a_0 + a_1 t_j + a_2 d_i. \tag{9}$$

The random effects parameters are assumed to follow a Gaussian distribution with a mean value of zero and a variance-covariance matrix that does not depend on individual i. Software written by Stram was used to provide restricted maximum likelihood estimation of model parameters (Laird et al., 1987). Other statistical programs such as BMDP(5V) (BMDP, 1990) and SAS PROC MIXED (SAS, 1991) can also be used to fit this model.

Generalized Estimating Equations (GEE) Method

The likelihood inference under the random effects modeling approach may be problematic due to a large number of parameters to be estimated. In addition, distributional assumptions required in this approach may not be always justified. The generalized estimating equations (GEE) method provides consistent estimates of the regression parameters and their variances under mild assumptions about their time dependence using the quasi-likelihood theory (Zeger and Liang, 1986). The estimating equations are derived without specifying the distributions of individual response measures. A brief description of this approach is given below.

Let (y_{ij}, x_{ij}), denote observations for times t_{ij} for $j = 1,2,...,n_i$ and $i = 1,2,...,N$ individuals. Here, y_{ij} represents the response variable and x_{ij} denotes a $p \times 1$ vector of covariates. Let y_i denote the vector $(y_{i1}, y_{i2},..., y_{ini})'$, $\mu_i = (\mu_{i1}, \mu_{i2},..., \mu_{ini})'$ the corresponding mean vector and x_i the $n_i \times p$ matrix $(x_{i1}, x_{i2},..., x_{ini})'$ for the ith individual and β a $p \times 1$ vector of regression coefficients.

The estimating equations are given by

$$\Sigma\, D_i' V_i^{-1} S_i = 0. \tag{10}$$

Here $S_i = y_i - \mu_i$, and $D_i = \partial\mu_i/\partial\beta$ and V_i is the "Working" Covariance Matrix for Y_i obtained by assuming the correlations among observations on the same individual as nuisance. Consistent estimate

of β and its asymptotic distribution was obtained under mild regularity conditions and using the actual covariance of y_i (Zeger and Liang, 1986).

We assume the mean (population) model given in equation 9. The software prepared by Karim and Zeger, (1988) which uses the Iteratively Reweighted Least Squares (IRLS) algorithm for parameter estimation was used.

RESULTS

Results from the analysis of the rabbit inhalation sperm data are summarized in this section. Total sperm count in the ejaculate was used as the biomarker of male fecundity. Table I-3 summarizes of the number of animals from which sperm samples were obtained in each dose group at each week of the DBCP inhalation study. In the 10 ppm dose group the number of animals decreases to half the starting number by exposure week 10 because of morbidity and mortality. Data from rabbit 34 in the 10 ppm dose group were excluded as the sperm counts recorded from week to week appeared inconsistent. Observations are missing due to death, sacrifice and insufficient sample or low sperm concentration.

Using the observations of semen volume and corresponding sperm concentration, the total sperm count (millions) in the ejaculate was calculated. Total sperm count in the ejaculate is considered by some investigators to be a better measure of male fecundity than sperm concentration. In Table I-4 and Figure I-2 means and standard deviations of the total sperm counts are presented for each week of the study. A substantial amount of variability is seen in the total sperm count measures for individual animals over time, between animals within each dose group, and across dose groups. Following the cessation of exposure, increasing trends are seen in the sperm count for animals other than those exposed to the 10 ppm dose group.

Individual animal normalized changes are calculated for each week by defining a baseline value for each animal from the two pre-exposure measures. The mean normalized changes for the four dose groups are presented in Table I-5 and Figure I-3. The number of negative values for the control group shows that the total sperm count are mostly lower than the pre-exposure values. This may suggest that when a repeated measure like sperm count, which can be modified by sampling, is used as a toxicological endpoint, the initial sampling period (control period) should extend for a sufficient period of time to allow the measure or endpoint to stabilize.

Table I-4. Total Sperm Count (million)[a,b] per Ejaculate in Control and DBCP Exposed Rabbits.

Week	ppm DBCP							
	0		0.1		1.0		10.0	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Pre-exposure								
−2	534.3	228.7	342.6	208.9	357.4	348.2	363.4	293.1
−1	278.5	121.7	321.7	145.7	229.9	117.1	215.0	115.2
Exposure								
1	594.9	384.6	525.1	365.3	277.5	188.6	432.1	279.2
2	492.6	261.8	480.2	235.6	328.0	215.5	192.0	116.4
3	572.1	350.4	605.4	313.7	462.6	252.5	336.3	235.8
4	414.4	232.0	631.1	408.4	351.5	144.6	430.6	127.9
5	300.3	194.4	350.2	224.2	316.3	204.9	427.4	212.3
6	666.6	164.2	451.8	310.4	530.7	265.1	399.4	283.7
7	534.9	357.6	335.1	187.0	337.3	151.2	231.7	183.4
8	318.6	122.5	485.3	397.4	263.6	147.0	69.7	123.9
9	440.0	158.4	569.4	203.3	358.5	141.4	61.7	137.7
10	605.0	326.7	441.9	263.7	261.4	174.8	7.5	9.1
11	474.4	256.5	415.9	309.0	182.0	163.3	3.9	2.7
12	805.8	664.0	496.7	236.5	231.1	141.0	0.1	0
13	647.2	414.6	811.8	661.8	173.7	138.3	3.5	6.4
14	466.2	326.3	523.9	340.1	73.6	116.6	1.1	1.2
Post-exposure								
16	515.9	251.3	558.6	374.9	214.8	168.3	1.1	0.9
19	436.2	390.8	609.7	415.4	209.6	293.9	0.3	0.2
24	531.2	572.2	760.6	474.0	411.3	308.1	0.3	0.7
26	390.0	154.5	750.9	199.8	246.7	92.6	2.7	5.7
27	512.3	354.1	511.3	319.5	216.8	57.9	9.6	14.1
28	258.2	187.3	414.4	433.1	352.0	118.4	6.4	12.6
30	627.0	737.4	447.7	293.5	308.5	368.7	4.4	6.0
32	235.6	151.4	623.4	208.7	282.7	173.5	7.7	9.9
34	419.9	469.0	331.2	112.6	488.2	263.1	7.2	12.0
36	288.4	205.4	493.6	280.4	341.5	183.4	2.3	3.2
38	829.5	720.7	492.5	350.7	291.3	286.3	15.6	20.2
40	451.9	333.4	733.8	386.5	309.0	259.1	34.7	64.5
42	288.0	229.7	272.2	234.0	293.8	203.1	1.8	2.4
44	458.8	278.5	485.3	279.2	369.9	271.9	8.1	8.5
46	447.2	380.5	615.8	306.7	324.4	117.6	36.9	57.1

[a,b] See Table I-3.

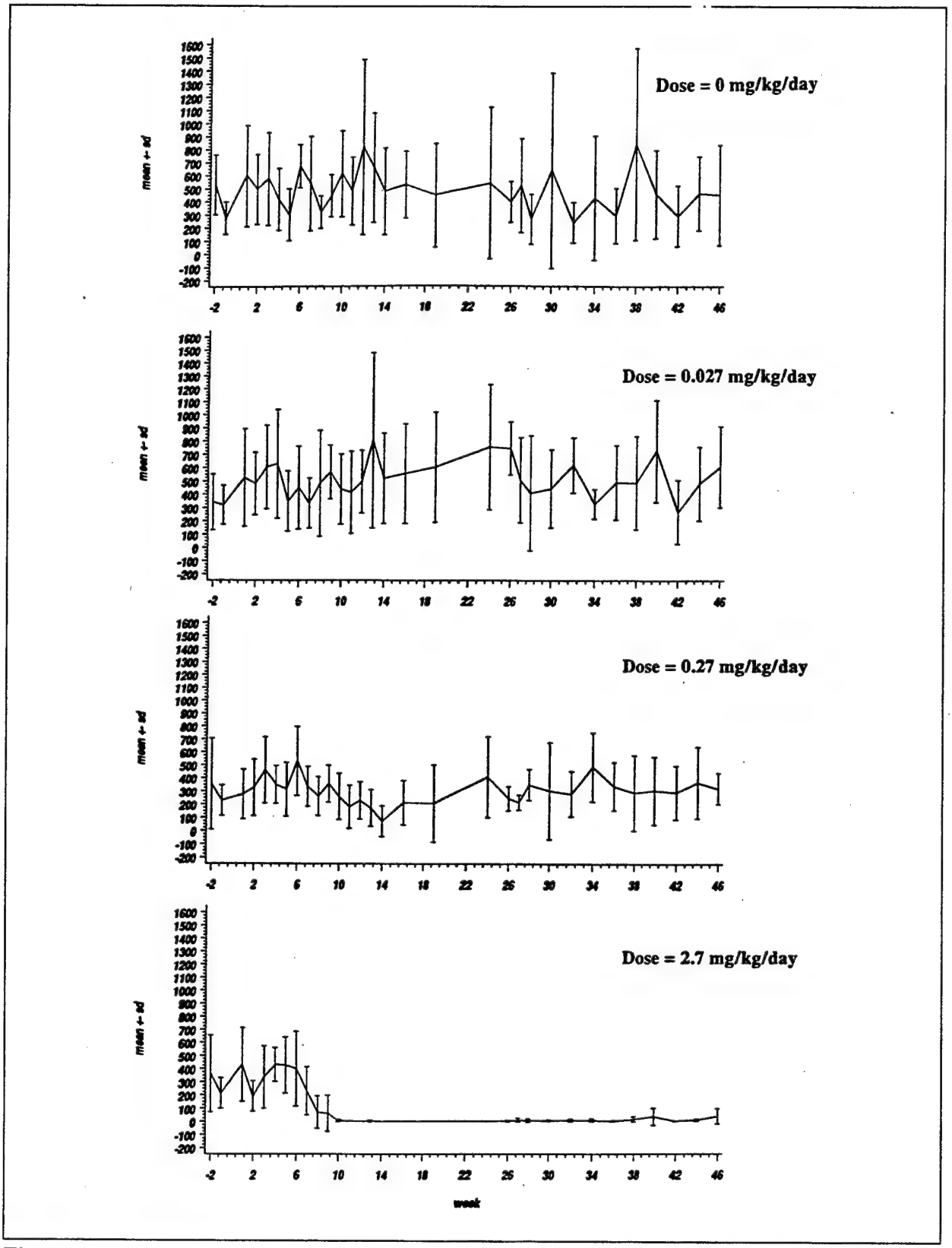


Figure I-2. Means ± SD of Total Sperm Count (millions) of Control and DBCP-Treated Animals.

Table I-5. Mean Normalized Changes[a] and Standard Deviations (SD) for Total Sperm Count (millions).

Week	ppm DBCP							
	0		0.1		1.0		10.0	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Exposure								
1	0.56	1.26	0.58	0.84	0.4	1.37	2.18	4.79
2	0.37	0.81	0.35	0.79	0.84	[illegible]9	– 0.159	0.65
3	0.49	1.06	0.97	1.09	1.38	1.43	1.3	2.09
4	0.06	0.58	1	1.36	0.77	1.15	1.27	1.51
5	–0.08	0.79	0.18	0.82	0.42	1.08	1.46	2.11
6	0.88	0.78	0.53	1.04	1.32	1.56	0.83	1.2
7	0.46	1.03	0.21	0.81	0.45	1	0.14	0.86
8	–0.11	0.51	0.4	1.04	0.13	0.72	–0.57	0.9
9	0.04	0.48	0.55	0.39	0.96	1.40	–0.58	0.98
10	0.57	1.06	0.48	0.92	0	0.49	–0.97	0.05
11	0.23	0.8	0.41	0.92	–0.2	0.83	–0.98	0.02
12	1.26	2.1	0.71	1.12	0.07	0.83	–1	0
13	0.83	1.35	1.67	1.81	–0.27	0.67	–0.97	0.06
14	0.27	1.02	1.22	1.96	–0.66	0.65	–0.99	0.01
Post-exposure								
16	0.4	0.99	1.46	1.96	0.04	1.10	–0.99	0.01
19	0.14	1.06	1.27	1.12	–0.18	[illegible]03	–1	0
24	0.12	0.62	2.12	1.96	0.85	[illegible]30	–1	0
26	0.01	0.57	1.79	0.93	–0.24	0.38	–0.98	0.03
27	0.22	0.77	0.92	0.99	–0.29	0.36	–0.89	0.2
28	–0.26	0.78	0.45	1.18	0.14	0.75	–0.96	0.07
30	0.19	0.77	0.72	0.70	–0.11	0.62	–0.96	0.04
32	–0.32	0.58	1.36	0.99	–0.09	1.03	–0.93	0.09
34	–0.14	0.59	0.44	0.55	0.67	0.35	–0.95	0.07
36	–0.19	0.8	1.08	1.32	0.04	0.62	–0.98	0.02
38	0.81	1.31	0.95	1.05	–0.33	0.66	–0.94	0.04
40	0.09	0.8	2.02	1.58	–0.14	0.54	–0.78	0.36
42	–0.38	0.33	–0.07	0.71	–0.19	0.69	–0.99	0.02
44	0.07	0.54	0.85	0.70	–0.02	0.65	–0.96	0.04
46	–0.11	0.26	1.51	1.35	0.19	0.84	–0.77	0.32

[a] Normalized change = (response-baseline)/baseline; baseline is defined as the average of the total sperm counts in the two pre-exposure weeks; a single entry is taken for the baseline if the other is missing; the group average is taken for the baseline when both entries are missing.

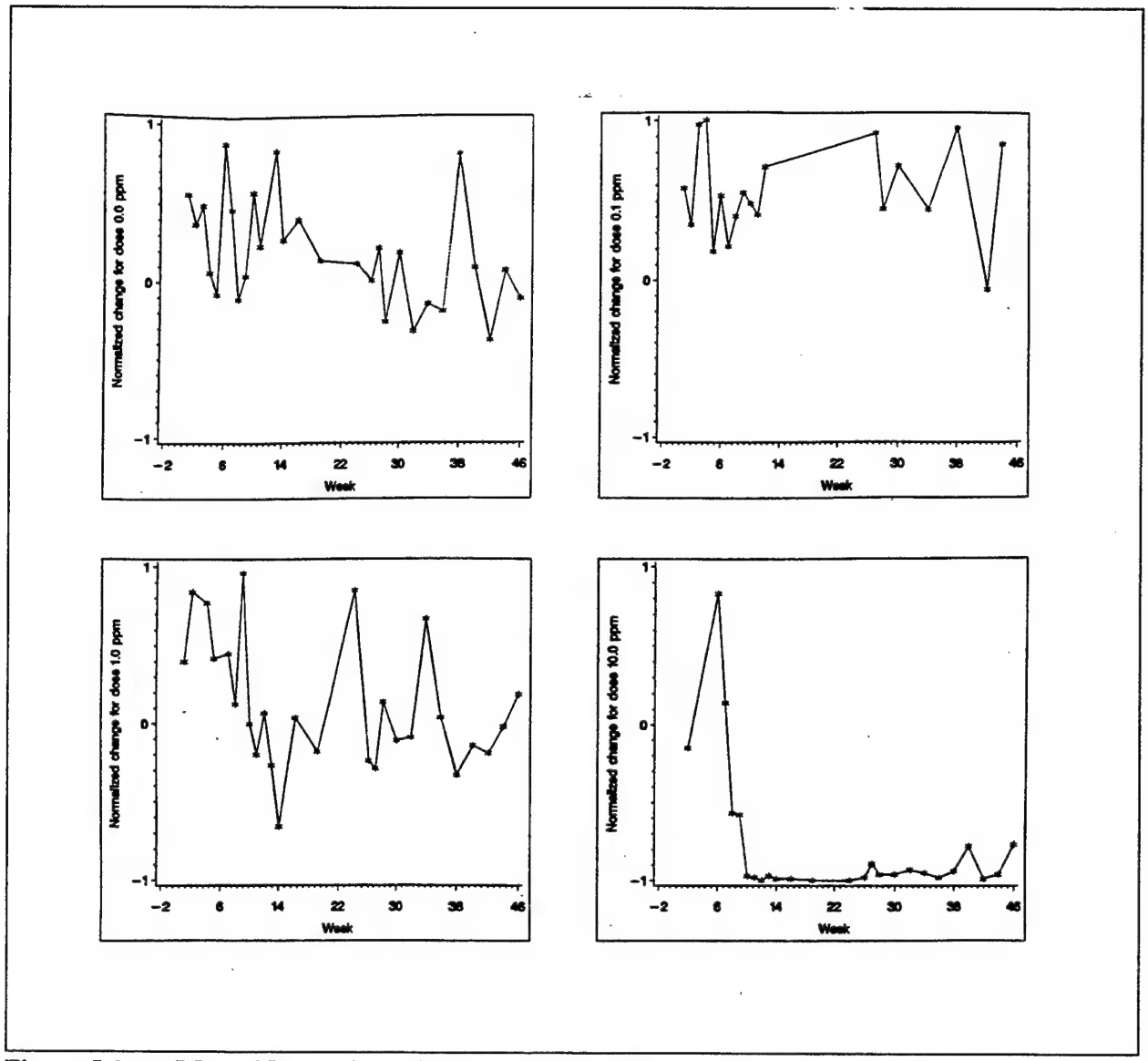


Figure I-3. Mean Normalized Changes in Total Sperm Count (millions) of Control and DBCP-Treated Rabbits (for then definition of normalized change see text).

Regression approach: Regression analyses were performed by using logarithms of the total sperm count values. Figure I-4 present results from weekly regression analyses using all dose groups. A sharp decrease in the weekly rates of change of log total sperm count is seen starting from week 7 and continuing through week 19. Thereafter, the drop is relatively less until the 46th week.

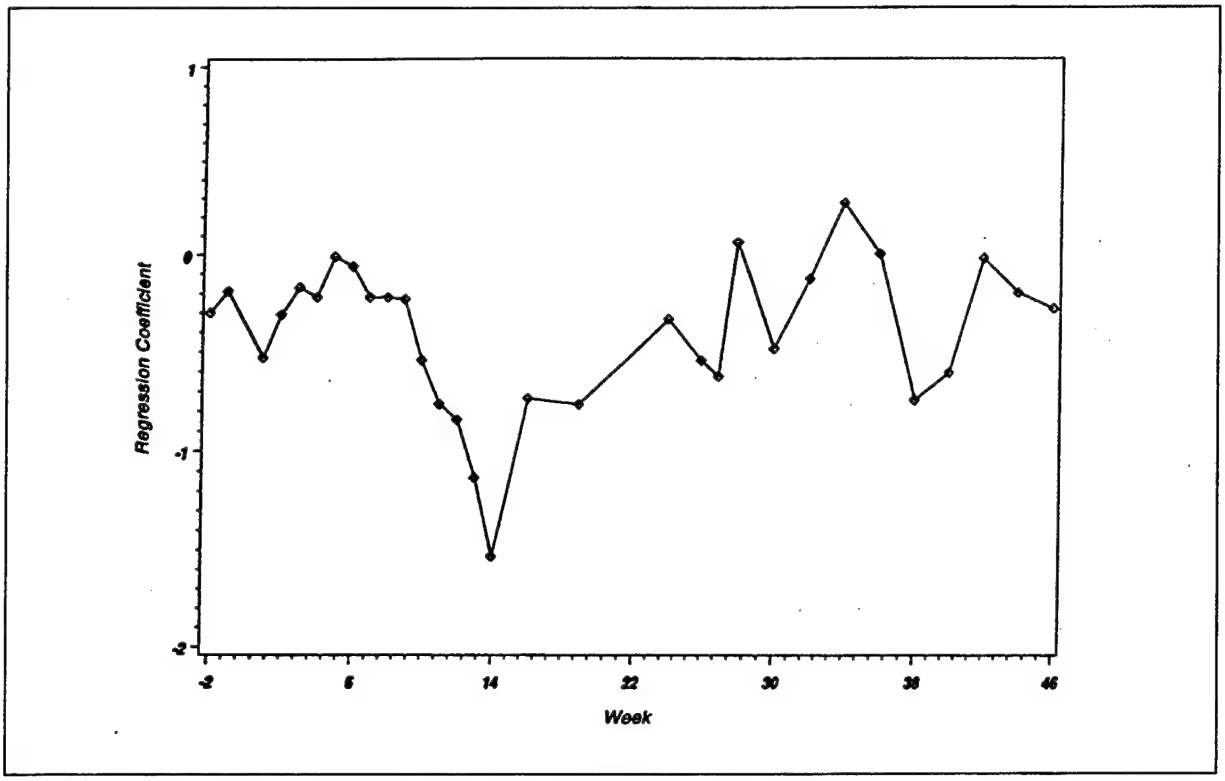


Figure I-4. Slope Estimates (log million/mg/kg/day) from Weekly Regressions of Total Sperm Count and DBCP Treatment (1 ppm for 6 h/day = 0.27 mg/kg/day).

Exploratory data analysis for the AVAS transformation: Using S-PLUS for additive and variance stabilizing transformations, a plot of the AVAS transformed values against the original values was created. The plot suggested a quadratic transformation (Figure I-5). This parametric form was fitted to the sperm number data using the AUTOREG procedure (SAS, 1990) to allow for possible correlated errors. The resulting equation is:

$$Y^* = -1.76432829 + 0.00439055 \times (\text{total sperm count}) - 0.000001042 \times (\text{total sperm count})^2 \quad (11)$$

where Y^* is called optimally transformed variable, with $R^2 = 79.0\%$. This means that 79.0% of the total variation of the AVAS transformation total sperm count is explained by a quadratic fit given by the above equation. The plot of the fitted (optimally transformed) values against the original values shows a quadratic pattern demonstrating the closeness with a quadratic form for the transformed values. Also, the plot of the residual values against the original sperm counts shows an even spread around the zero value on the vertical axis, indicating randomness of the errors (Figure I-5). Next, Equation 11 was used to transform the total sperm count values in all other dose groups. The parametric transformation based on AVAS transformed variables is referred to as the "optimal" transformation.

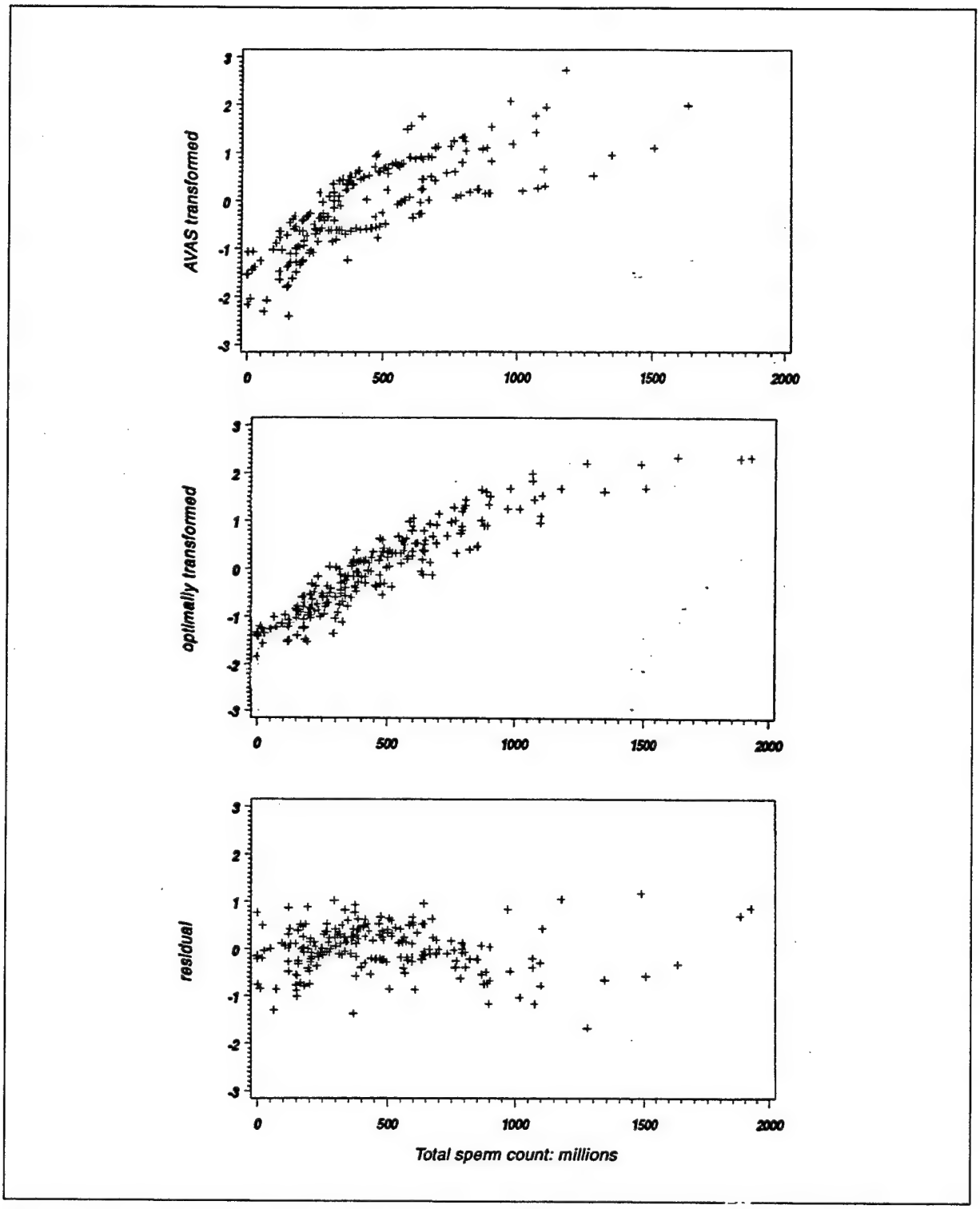


Figure I-5. **AVAS-Transformed, Optimally Transformed, and Residual Total Sperm Count of Control Rabbits Plotted Against the Original Total Sperm Count Values.**

Dose–Response Modeling: In addition to using the AVAS-based transformed measures, two other commonly used transformations are employed in the present analysis. These are logarithmic and cube root transformations. We note that as the logarithmic transformation is a special case of Box-Cox transformations, we did not use explicitly the Box-Cox methodology. The AVAS based transformation is used here to illustrate the potential use of exploratory non-parametric data transformation techniques in this area. Results from model fitting using REML estimation for the random-effects models and GEE method and three transformed data sets pertaining to the entire duration (week-2 through 46) and a partial duration (week-2 through week-19) are presented in Table I-6. The time and dose estimates are negative indicating that with increasing dose, the total sperm count decreases and the effect of DBCP exposure increases over time with continued exposure. The time coefficients are found to be relatively smaller when data from the total time period are used compared to the situation when data through week 19 are used. A reverse pattern is seen in the dose coefficients.

Table I-6. Regression Estimates[a] (standard errors) from Logitudinal Dose–Response Models of Total Sperm Count and DBCP Exposure Using Restricted Maximum Likelihood (REML) Method for Random Effects Models, Generalized Estimating Equations (GEE) Methods and Three Variance Stabilizing Transformations.

Transformation	Period Analyzed							
	<-2 - 19> Weeks				<-2 -46> Weeks			
	REML		GEE		REML		GEE	
Optimal[b]								
Intercept	-0.10	(0.08)	0.01	(0.10)	-0.10	(0.05)	0.03	(0.11)
Time	-0.02	(0.01)	-0.01	(0.01)	-0.01	(0.01)	-0.02	(0.01)
Dose	-0.23	(0.09)	-0.34	(0.07)	-0.56	(0.13)	-0.28	(0.06)
Logarithmic								
Intercept	6.08	(0.10)	6.79	(0.30)	6.01	(0.10)	6.44	(0.24)
Time	-0.10	(0.03)	-0.11	(0.04)	-0.04	(0.01)	-0.05	(0.02)
Dose	-0.23	(0.18)	-0.74	(0.15)	-0.85	(0.20)	-0.93	(0.19)
Cube Root								
Intercept	7.43	(0.16)	8.08	(0.32)	7.22	(0.12)	7.81	(0.27)
Time	-0.09	(0.03)	-0.10	(0.04)	-0.03	(0.02)	-0.05	(0.02)
Dose	-0.31	(0.15)	-0.17	(0.01)	-0.01	(0.01)	-1.08	(0.21)

[a] Estimates are based on the conversion formula: 1.0 ppm for 6 hours per day is equivalent to 0.27 mg/kg/day.

[b] Optimally transformed total sperm count = $-1.76432829 + 0.00439055$ (total sperm count) $- 0.000001042$(total sperm count)^2

Benchmark Dose Calculation

An important purpose of the present investigation is to evaluate a derived dose–response relationship in the context of DBCP exposure of rabbits and suggest a BD which is defined as the $L_{95}ED_{10}$ (Crump, 1984; Pease et al., 1991). We have obtained regression estimates for the dose coefficient from a longitudinal linear model using two different inferential procedures—REML and GEE with three different types of transformation, optimal (AVAS-based) logarithmic and cube root. Considering a baseline value of 400 million for the total sperm count these estimates are presented in Table I-7.

DISCUSSION AND CONCLUSIONS

The principal advantages of using the statistical approaches based on longitudinal data analysis techniques, described in this paper to subchronic toxicological study data, are to take into account the time dependence of the response measures in dose–response modeling and to take both within and between individual variabilities into account. An exploratory data analysis technique is used in the present study to derive a suitable transformation for the sperm measures to reduce the variability. This procedure reduces the variations in the endpoints by modeling the error component as additive to the systematic variation and using the asymptotic variance stabilizing transformation. In addition, commonly used transformations such as logarithmic (Pease et al., 1991) and cube root (Meistrich and Brown, 1983) are also evaluated.

By applying these techniques to different segments of the study period, it is apparent that the benchmark dose estimates differ. This is an important finding and is partly due to the fact that the cycle of spermatogenesis is altered in the presence of exposure to DBCP and the sperm measures are highly variable in the control animals. In the 10 ppm dose group, about half the animals are lost by either sacrifice or death due to the exposure by week 8. The remaining animals in this group show at best little or no recovery in sperm counts after cessation of exposure. By taking different segments of the experiment, a null hypothesis that the slopes of the dose on sperm measures are equal across weeks was tested. By rejecting this hypothesis, we note that the effect of dose on sperm count in the ejaculate varies from week to week within the two different segments considered. A risk averse approach to reproductive risk assessment would identify the week with the optimal (or largest negative) dose parameter for defining or estimating the BD. However, that approach does not take advantage of the total information provided by the longitudinal study. In addition, these analyses suggest that chronic experiments should be continued until the estimate of the dose parameter stabilizes to measure the

"true" effect of the toxicant. For example, the BD estimate from these analyses may be substantially greater than the actual BD calculated if the study were carried to the point where the BD stabilized or the dose parameter stabilized. In that sense even the "risk-averse" single parameter estimate of the BD may actually under estimate the effect.

When characterizing risk only from an exposure period, the modeling process does not take into account the recovery of spermatogenesis after cessation of exposure. The effect of dose and time are different in this period when compared to the entire study period including a recovery period. For endpoints that demonstrate reversible damage to testicular function, it is possible that an adverse effect from a given dose persists only for a limited time after cessation of exposure (an obvious exception is the highest dose level at which complete irreversible disruption of the seminiferous epithelium occurs). This suggests that when establishing allowable levels of exposure to toxicants like DBCP, it is useful to consider the entire study period rather than one specific segment. Also, time and dose maintain an inverse relationship with the endpoint considered in this analysis. This means that with increasing dose the sperm count decreases and the longer the exposure period, the greater the dose effect on sperm count. Indeed, in this experiment in which exposure was conducted for the duration of the cycle of spermatogenesis, the maximum dose effect was not determined. It may also be mentioned that another dose metric such as the cumulative dose will be interesting to explore in the above context for establishing the dose x time interactions and obtaining a BD estimate in such studies.

Results obtained from the analyses in this paper are exploratory in nature and form the basis for further investigation of statistical methods for time dependent dose–response relationships in reproductive toxicology. It is also believed that in view of excessive variations among the study animals, statistical analyses are requisite to address questions of inter- and intra-animal variability. Statistical models of longitudinal data have usually two objectives. First, they allow us to cast the problem in a regression framework which relates the response variable to a set of covariates and secondly, they address ways and means to account for the within individual correlation arising from repeated observations. We can also incorporate for individual animals their corresponding weekly body weights, organ weights and/or other relevant biological parameters as covariates that may show a causal link with the response to any possible extent.

A general model has been presented integrating elements necessary for assessment and quantitation of reproductive risks following male exposure. Once the relationship between exposure and male fecundity, female fecundity, and couple fecundity has been defined, it is possible to explore the

relationship to reproductive toxicity. The biomarker based model presented is a first step in the development of quantitative approaches for estimation of male mediated reproductive risks. Eventually it is hoped that more complete definition of MF, FF, CF could all be combined into one model for reproductive risk assessment.

Table I-7. Estimates[a] of Benchmark Dose for DBCP Using the Total Sperm Count.

Transformation	$L_{95}ED_{10}$ (mg/kg/day)[a]			
	<-2 - 19> Weeks		<-2 - 46> Weeks	
	REML[b]	GEE[c]	REML	GEE
Optimal[d]	0.355	0.302	0.177	0.362
Logarithmic	0.172	0.097	0.081	0.077
Cube root	0.475	0.239	8.620	0.192

[a] Estimates were based on the conversion formula: 1 ppm for 6 hours is equivalent to 0.27 mg/kg/day. $L_{95}ED_{10}$ is defined as the 95% lower confidence limit on the dose level producing a 10% change in response from the background response which is taken as 400 million total sperm count.

[b] REML: Restricted Maximum Likelihood estimation for random-effects models.

[c] GEE: Generalized Estimating Equations Method.

[d] Optimally transformed total sperm count = $-1.76432829 + 0.00439055$(total sperm count) $- 0.000001042$(total sperm count)^2

REFERENCES

Aafjes, J.H., Vels, J.M., and Schenck, E. 1980. Fertility of rats with artificial oligozoospermia. *J. Reprod. Fertil.* 58:345–351.

Amann, R.P. and Hammerstedt, R.H. 1980. Validation of a system for computerized measurements of spermatozoal velocity and percentage of motile sperm. *Biol. Reprod.* 23:647–656.

Amann, R.P. and Howards, S.S. 1980. Daily spermatozoal production and epididymal spermatozoal reserves of the human male. *J. Urol.* 124:211–215.

Baird, D.D., Wilcox, A.J., and Weinberg, C.R. 1986. Use of time to pregnancy to study environmental exposures. *Am. J. Epidemiol.* 124: 470–480.

Baird, D.D. and Wilcox, A.J. 1985. Cigarette smoking associated with delayed conception. *JAMA* 253:2979–2983.

Barnes, D.G. and Dourson, M. 1988. Reference dose (RfD): description and use in health risk assessments. *Regul. Toxicol. Pharmacol.* 8:471–486.

Barrett, J.C. and Marshall, J. 1969. The risk of conception on different days of the menstrual cycle. *Pop. Studies* 23:455–461.

Barrett, J.C. 1971. Fecundability and coital frequency. *Pop. Studies* 25:309–313.

BMPD Statistical Software Manual, 1990. University of California Press, 2120 Berkeley Way, Berkeley, CA 94720.

CECOS, Schwartz, D. and Mayaux M.J. 1982. Female fecundity as a function of age — Results of artificial insemination in 2193 nulliparous women with azoospermic husbands. *N. Engl. J. Med.* 306:404–406.

Crump, K.S. 1984. A new method for determining allowable daily intakes. *Fundam. Appl. Toxicol.* 4:854–871.

David, G., Jouannet, P., Martin-Boyce, A., Spira, A., and Schwartz, D. 1979. Sperm counts in fertile and infertile men. *Fertil. Steril.* 31:453–455.

Dobbins, J.G. 1987. Regulation and the use of "negative" results from human reproductive studies: the case of ethylene dibromide. *Am. J. Ind. Med.* 12:33–45.

Eaton, M., Schenker, M., Whorton, D., Samuels, S., Perkins, C., and Overstreet, J. 1986. Seven-year follow-up of workers exposed to 1,2-dibromo-3-chloropropane. *J. Occup. Med.* 28:1145–1150.

Ewing, L.L. and Mattison, D.R. 1987. Introduction: biological markers of male reproductive toxicology. *Environ. Health Perspect.* 74:11–13.

Faustman, E.M., Wellington, D.G., Smith, W.P., and Kimmel, C.A. 1989. Characterization of a developmental toxicity dose-response model. *Environ. Health Perspect.* 229–241.

Foote, R.H., Schermerhorn, E.C., and Simkin, M.E. 1986a. Measurement of semen quality, fertility, and reproductive hormones to assess dibromochloropropane (DBCP) effects in live rabbits. *Fundam. Appl. Toxicol.* 628–637.

Foote, R.H., Berndtson, W.E., and Rounsaville, T.R. 1986b. Use of quantitative testicular histology to assess the effect of dibromochloropropane (DBCP) on reproduction in rabbits. *Fundam. Appl. Toxicol.* 638–647.

Generoso, W.M. 1980. Repair in fertilized eggs of mice and its role in the production of chromosomal aberrations. *Basic Life Sci.* 15:411–420.

Generoso, W.M., Huff, S.W., and Cain, K.T. 1979a. Relative rates at which dominant-lethal mutations and heritable translocations are induced by alkylating chemicals in postmeiotic male germ cells of mice. *Genetics* 93:163–171.

Generoso, W.M., Cain, K.T., Krishna, M., and Huff, S.W. 1979b. Genetic lesions induced by chemicals in spermatozoa and spermatids of mice are repaired in the egg. *Proc. Natl. Acad. Sci. USA* 76: 435–437.

Gladen, B.C., Williams, J. and Chapin, R.E. 1991. Issues in the Statistical Analysis of Sperm Motion Data Derived from Computer-assisted Systems. *Journal of Andrology*, 12(2):89–97.

Goldsmith, J.R., Potashnik, G., and Israeli, R. 1984. Reproductive outcomes in families of DBCP-exposed men. *Arch. Environ. Health* 39(2):85–89.

Guerrero, V.R. and Rojas, O.I. 1975. Spontaneous abortion and aging of human ova and spermatozoa. *N. Engl. J. Med.* 293:573–575.

Horning S.J., Hoppe R.T., Kaplan H.S., and Rosenberg S.A., 1981. Female reproductive potential after treatment for Hodgkin's disease. *N. Engl. J. Med.* 1377–1382.

Jasanoff S. 1990. The Fifth Branch. Science Advisors as Policymakers. *Harvard University Press,* Cambridge, MA.

Karim, M.R. and Zeger, S.L. 1988. GEE: A SAS Macro for longitudinal data analysis, Version 1, *Technical Report No. 674, Dept. of Biostatistics, The Johns Hopkins University, Baltimore, MD.*

Krewski, D. and Franklin, C. (editors). 1991. Statistics in Toxicology. *Gordon and Breach Science Publishers, New York.*

Laird, N.M. and Ware, J.H. 1982. Random-Effects Models for Longitudinal Data. *Biometrics* 38:963–974.

Laird, N., Lange, N., and Stram, D. 1987. Maximum Likelihood Computations With Repeated Measures: Application of the EM Algorithm. *JASA* 82(397):97–105.

Lamb, J.C. 1989. Design and use of multigeneration breeding studies for identification of reproductive toxicants. *Toxicology of the Male and Female Reproductive Systems, Working, P.K. (editor), Hemisphere Publishing Corporation, New York:*131–155.

Levine, R.J., Symons, M.J., Balogh, S.A., Arndt, D.M., Kaswandik, N.T., and Gentile, J.W. 1980. A method for monitoring the fertility of workers. 1. Method and Pilot studies. *J. Occup. Med.* 22:781–791.

Levine, R.J., Symons, M.J., Balogh, S.A., Milby, T.H., and Whorton, M.D. 1981. A method for monitoring the fertility of workers. 2. Validation of the method among workers exposed to dibromochloropropane. *J. Occup. Med. 23.* 183–188.

Mattison, D.R., Working, P.K., Blazak, W.F., Hughes, C.L., Killinger. J.M., Olive, D.L., and Rao, K. S. 1989. Criteria for identifying and listing substances known to cause reproductive toxicity under California's Proposition 65, *Reprod. Toxicol.* 4:163–175.

Mattison, D.R. 1991. An overview on biological markers in reproductive and developmental toxicology: concepts, definitions and use in risk assessment. *Biomed. Environ. Sci.* 4:8–34.

Mattison, D.R, Plowchalk, D.R., Meadows, M.J., Al-Juburi, A.Z., Gandy, J., and Malek, A. 1990. Reproductive toxicity: the male and female reproductive systems as targets for chemical injury. *Environ. Med.,* 74:391–411.

Mattison, D.R. and Thomford, P.J. 1989. The mechanisms of action of reproductive toxicants. *Toxicol. Pathol.* 17:364–376.

Mattison, D.R., Blann, E., and Malek, A. 1991. Physiological alterations during pregnancy: impact on toxicokinetics. *Fundam. Appl. Toxicol.* 16:215–218.

Mattison, D.R and Brewer, D.W. 1988. Computer modeling of human fertility: the impact of reproductive heterogeneity on measures of fertility. *Reprod. Toxicol.* 2:253–271.

Mattison, D.R. 1990. Fetal pharmacokinetic and physiological models. *Reproductive Toxicology: Risk Assessment and the Future, Hood, R.D. (editor), Van Nostrand Reinhold, New York*: 137– 154.

Mattison, D.R. 1993. Sites of female reproductive vulnerability: implications for testing and risk assessment. *Reprod. Toxicol.* 7:53–62.

Meistrich, M.L. 1989a. Interspecies comparison and quantitative extrapolation of toxicity to the human male reproductive system. *Toxicology of the Male and Female Reproductive Systems, Working,* P.K. (editor), Hemisphere Publishing Company, New York: 303–321.

Meistrich, M.L. 1989b. Calculation of the incidence of infertility in human populations from sperm measures using the two-distribution model. *Sperm Measures and Reproductive Success: Institute for Health Policy Analysis Forum on Science, Health, and Environmental Risk Assessment, Burger, E.J., Tardiff, R.G., Scialli, A.R., Zenick, H. (editors), Alan R. Liss, New York*: 275–290.

Meistrich, M.L. and Brown, C.C. 1983. Estimation of the increased risk of human infertility from alterations in semen characteristics. *Fertil. Steril.* 40:220–230.

Meistrich, M.L. 1992. A method for quantitative assessment of reproductive risks to the human male. *Fundam. Appl. Toxicol.* 18:479–490.

Meistrich, M.L. 1984. Stage-specific sensitivity of spermatogonia to different chemotherapeutic drugs. *Biomed. Pharmacother.* 38:137–142.

Meistrich, M.L. 1988. Estimation of human reproductive risk from animal studies: determination of interspecies extrapolation factors for steroid hormone effects on the male. *Risk Anal.* 8:27–33.

Menken, J., Trussell, J., and Larsen, U. 1986. Age and infertility. *Science* 233:1389–1394.

Milby, T.H. and Whorton, D. 1980. Epidemiological assessment of occupationally related, chemically induced sperm count suppression. *J. Occup. Med.* 22:77–82.

Morrissey, R.E. 1989. Association of sperm, vaginal cytology, and reproductive organ weight data with fertility of Swiss (CD-1) mice. *Toxicology of the Male and Female Reproductive Systems,* Working, P.K. (editor). Hemisphere Publishing Corporation, New York. 199–216.

Myers, R.H. 1990. *Classical and Modern Regression With Applications*, 2nd edition. PWS–Kent Publishing Company, Boston.

National Research Council 1989. *Biologic Markers in Reproductive Toxicology.* National Academy Press, Washington, DC.

Paul, M. (editor) 1993. *Occupational and Environmental Reproductive Hazards: A guide for clinicians*, Williams and Wilkins, Baltimore.

Pease, W., Vandenberg, J., and Hooper, K. 1991. Comparing alternative approaches to establishing regulatory levels for reproductive toxicants: DBCP as a case study. *Environ. Health Perspect.*, 91:141–155.

Potashnik, G. 1983. A four-year reassessment of workers with dibromochloropropane-induced testicular dysfunction. *Andrologia* 15: 164–170.

Potashnik, G. and Abeliovich, D. 1985. Chromosomal analysis and health status of children conceived to men during or following dibromochloropropane-induced spermatogenic suppression. *Andrologia* 17:291–296.

Rao, K.S., Burek, J.D., Murray, F.J., John, J.A., Schwetz, B.A., Beyer, J.E., and Parker, C.M. 1982. Toxicologic and reproductive effects of inhaled 1, 2-dibromo-3-chloropropane in male rabbits. *Fundam. Appl. Toxicol.* 2:241–251.

Rao, K.S., Burek, J.D., Murray, J.S., John, J.A., Schwetz, B.A., Murray, F.J., Crawford, A.A., Deacon, M.M., Potts, W.J., Sutter, B.N., Dittenber, D.A., Bell, T.J., Beyer, J.E., Albee, R.R., Battjes, J.E., and Parker, C.M. 1980. 1, 2-Dibromo-3-Chloropropane: Inhalation Fertility Study in Rats and Rabbits. Toxicological Research Laboratory, Health and Environmental Sciences USA and Dow Chemical USA, Midland, Michigan.

Ratcliffe, J.M., Schrader, S.M., Steenland K., Clapp, D. E., Turner, T., and Hornung, R.W. 1987. Semen quality in papaya workers with long term exposure to ethylene dibromide, *Br. J. Ind. Med.*, 44:317–326.

S-PLUS for DOS Reference Manual: Version 2.0. 1992. Statistical Sciences, Inc. Seattle, Washington.

SAS User's Guide, Version 6.03. 1991. SAS, Inc. Cary, NC.

Saegusa, J. 1989. Cumulative effects of 1, 2-dibromo-3-chloropropane (DBCP) on kidney and testis. *Indust. Health* 27:49–58.

Scialli, A.R. and Zinaman, M.J. (editors) 1993. *Reproductive Toxicology and Infertility*. McGraw-Hill, Inc.

Tibshirani, R. 1988. Estimating transformations for regression via additivity and variance stabilization. *JASA*. 83:394–405.

Ware, J.H., Dockery, D.W., Louis, T.A., Xu, X., Ferris, B.G., and Speizer, F.E. 1990. Logitudinal and cross-sectional estimates of pulmonary function decline in never-smoking adults. *Am. J. Epidemiol.* 132(4):685–700.

Whorton, D., Krauss, R.M., Marshall, S., and Milby, T.H. 1977. Infertility in male pesticide workers. *The Lancet*. 1259–1261.

Whorton, D., Milby, T.H., Krauss, R.M., and Stubbs, H.A. 1979. Testicular function in DBCP exposed pesticide workers. *J. Occup. Med.* 21:161–166.

Whorton, M.D., Milby, T.H., Stubbs, H.A., Avashia, B.H., and Hull, E.Q. 1979. Testicular functions among carbaryl-exposed employees. *J. Toxicol. Environ. Health* 5:929–941.

Williams, J., Gladen, B.C., Schrader, S.M., Turner, T.W., Phelps, J.L., and Chapin, R.E. 1990. Semen analysis and fertility assessment in rabbits: statistical power and design considerations for toxicology studies. *Fundam. Appl. Toxicol.* 15:651–665.

Williams, J., Gladen, B.C., Turner, T.W., Schrader, S.M., and Chapin, R.E. 1991. The effects of ethylene dibromide on semen quality and fertility in the rabbit: evaluation of a model for human seminal characteristics, *Fundam. Appl. Toxicol.* 16:687–700.

Working, P.K. and Mattison, D.R. 1993. Reproductive and Reproductive Toxicity Testing Methods in Animals. *Occupational and Environmental Reproductive Hazards,* Paul, M. (editor) Williams and Wilkins, Baltimore, MD: 91–99.

Zeger, S.L. and Liang, K-Y. 1986. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 42:121–130.

SESSION II

Temporal Factors

of Exposure in Identifying Hazards

CRITICAL PERIODS OF EXPOSURE AND DEVELOPMENTAL OUTCOME

George P. Daston
Miami Valley Laboratories, The Procter & Gamble Company, Cincinnati, OH 45239

Jeanne M. Manson
Dept. of Safety Assessment, Merck Research Laboratories, West Point, PA 19486

Development is characterized by an extremely rapid progression through a number of unique stages. It is well known that the organism progresses very quickly from a single cell to a complex being consisting of billions of cells. However, proliferation is arguably the least complicated of developmental processes. The embryo must also establish a body plan, partition function, and elicit the differentiation of diverse tissues and organs in the appropriate positions, at the appropriate times. The embryo simultaneously undergoes a series of dramatic physiological adaptations to cope with its changing status: initially as a very small ball of cells whose metabolic needs can be fulfilled via diffusion of nutrients in an oxygen-poor environment; then, as a larger, more complex embryo with a circulatory system whose nutritional and gas exchange needs are served by a placenta; to a free-living, air breathing creature at birth.

Clearly this rapid progression sets the developing organism apart from the mature one. It also makes the embryo, fetus, and child sensitive to toxic insults in a way that is unique to the stage of development at which the exposure occurs: the response to an insult is dependent on the developmental stage during exposure. There are specific temporal windows, known as critical periods of development, for each different manifestation of abnormal development. It is the purpose of this brief review to summarize our knowledge of those periods.

Figure II-1, taken from Wilson's *Environment and Birth Defects* (Wilson, 1973), illustrates that the peak sensitivity to teratogenesis is during the period of organ formation, roughly from 14 days-8 weeks post-conception in humans, 6–15 days post-conception in rats and mice. Prior to this period, Wilson hypothesized that intoxication would lead to embryolethality, but not malformation; after organogenesis, the fetus would be increasingly resistant to malformation and death, but toxicity would be manifested as growth retardation and functional deficit. Information accumulated over the past two decades have corroborated the validity of the scheme, with some modifications. The susceptibility of each developmental period will be evaluated below.

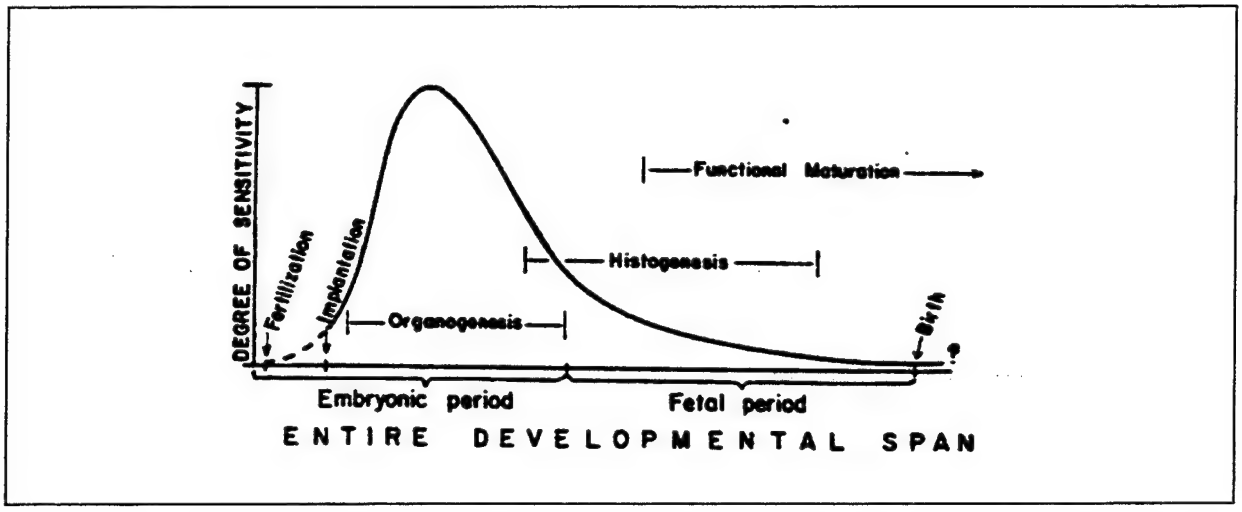


Figure II-1. Wilson's Scheme Demonstrating That Organogenesis is the Most Sensitive Period to Teratogenesis. Earlier stages are largely resistant to teratogenesis, but susceptible to embryolethality. The fetal and postnatal period are periods of growth and functional maturation, and insult during these periods would mainly affect these processes. From Wilson (1973), with permission.

Early Development (Fertilization to Gastrulation): The early stages of development are characterized by rapid cell division, but little differentiation of function. This stage has been characterized as refractory to teratogenesis, but sensitive to embryolethality. Rates of spontaneous abortion during this period are quite high. A rate of 35% has been determined using a sensitive immunoassay for detecting pregnancy (Wilcox et al., 1985), but may be somewhat higher as the assay is not 100% reliable during the first week of pregnancy. The relative contribution of environmental agents to this early loss is unknown. A large proportion of abortuses recovered are aneuploid or polyploid, suggesting faulty maturation or meiosis of germ cells rather than an effect on the early embryo.

It may be speculated that the refractoriness of this early period to teratogenesis is attributable to the fact that the developmental fate of embyonic cells has not been rigidly determined. Experiments with chimeric mice indicate that only a few cells [as few as three (Markert and Petters, 1978)] within the blastocyst give rise to the entire embryo. There does not appear to be any rigid determination within the inner cell mass as to which cells form the embryo, as cells introduced into the inner cell mass frequently become embryonic progenitors. (This plasticity is the basis for gene "knockout" technology.) Therefore, an adverse effect on cells of the cleavage and blastocyst-stage embryos is unlikely to have a teratogenic effect, as long as there remained unaffected cells to form the embryo. It is likely that exceptions will be found to this general rule, however.

One notable exception that dramatically illustrates the extreme transience of critical periods is the observation that a number of potent mutagens cause malformations when administered immediately after fertilization. The agents shown to do this include ethylene oxide (Generoso et al., 1987; Rutledge and Generoso, 1989), ethylnitrosourea, ethylmethane sulfonate, and triethylenemelamine (Generoso et al., 1988). Frank abnormalities were produced in fetuses after treatment 1 or 6 hours post-fertilization, while only marginal effects were produced after treatment 9 hours post-fertilization. It is unlikely that this exquisite sensitivity is due to mutations passed on to all the daughter cells, as the embryo is still one-celled until approximately 24 hours post-fertilization, which does not match the duration of the critical period for these effects.

Organogenesis: This period of development is extremely sensitive to malformations (Figure II-1). During this period cell fate becomes determined, and groups of cells undergo morphogenetic processes to create organ anlage, whose shape is further defined by additional morphogenetic events. Failure in any of these processes may lead to abnormal formation of the entire structure. Therefore, the critical period for structural malformations is just prior to and during organ formation. However, not all structures form at the same time or rate; therefore, each structure has its own critical period of susceptibility (Figure II-2). A teratogenic exposure early during organogenesis might be expected to affect development of the central nervous system, eyes or heart; a later treatment the kidneys or palate, and still later the genitalia.

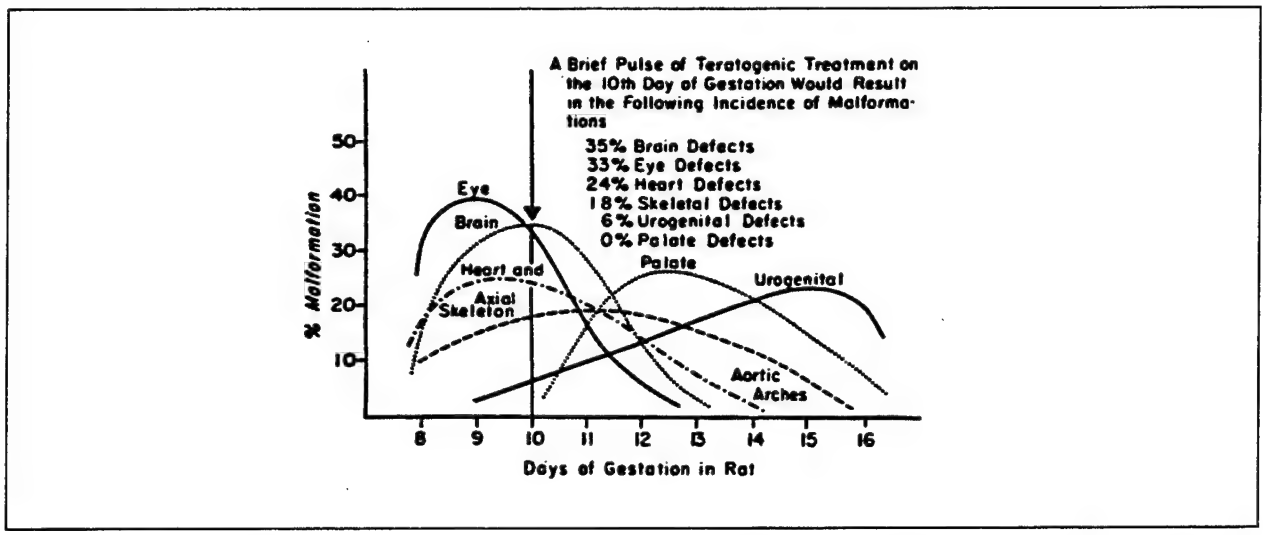


Figure II-2. Critical Periods of Sensitivity for Several Organs. This figure illustrates that the adverse outcome is dependent on the time of exposure. From Wilson (1973), with permission.

The importance of time of exposure to teratogenic outcome is well illustrated by data on ethylenethiourea (ETU), a fungicide breakdown product that produces a diverse array of malformations in rat embryos. Ruddick and Khera (1975) evaluated the defects produced after a single oral administration of a teratogenic dosage of ETU on gestation days 9–20 (sperm positive = day 0). As expected, the types and frequencies of malformations were dependent on the time of treatment. Eye defects were only seen after treatment on day 10 or 11. Cleft palate was produced with treatments on days 11–15. Defects of the digits were produced in both fore- and hindlimbs, but the frequency of these defects reached a peak earlier in the forelimbs, reflecting the fact that they develop before the hindlimbs. Table II-1 illustrates the critical periods for selected malformations after ETU treatment.

Table II-1. Frequency of Abnormalities in Rats Given a Single Oral Dosage of 240 mg/kg ETU on the Indicated Day of Gestation[1].

	Day Of Treatment[2]							
Abnormality	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Eye defects	¾	2	15	¾	¾	¾	¾	¾
Ectrosyndactyly								
-forelimb	¾	¾	100	75	100	40	¾	¾
-hindlimb	¾	¾	12	25	100	40	¾	¾
Tail defects	20	100	85	100	67	¾	¾	¾
Cleft palate	¾	¾	68	50	34	70	100	¾

[1] Data from Ruddick and Khera (1975)

[2] Gestation days calculated from sperm positive = day 0.

Once an organ has formed it becomes less susceptible to teratogenesis. However, there are instances where a structure can be malformed even after organogenesis. For example, the most serious central nervous system defects, neural tube defects, are generally thought to arise when the cells of the neural plate fail to form a closed tube that then develops into brain and spinal cord. However, neural tube defects have also been induced in the laboratory after neural tube closure. The teratogen ETU causes re-opening of the neural tube of rat embryos (Rogers et al., 1994). This appears to result from an excess accumulation of fluid within the tube due to faulty osmoregulation (Daston et al., 1987). The most recently discovered class of human teratogens, the angiotensin converting enzyme (ACE) inhibitors, also act after organogenesis. These agents appear to decrease renal blood flow in the fetus (Martin et al., 1992), which in turn leads to decreased urine production and oligohydramnios (decreased amniotic fluid volume). Oligohydramnios may contribute to skull hypoplasia because the pressure of the uterine musculature is not dispersed by the amniotic fluid (Brent and Beckman, 1991). An alternative explanation is that agent-induced fetal hypotension adversely affects calvarial development (Barr and Cohen, 1991). These two examples illustrate that malformations can occasionally be produced after the critical period for organ formation; however, for the most part teratogenesis is limited to this period.

The critical periods for organogenesis are typically quite brief, both in laboratory animals and humans (Table II-2). This is the basis for the rapidly changing sensitivity of the embryos, as well as the changing pattern of effects.

Table II-2. Important Developmental Events In Rat, Rabbit, Monkey, and Human.*

Event	Day Of Gestation			
	Rat	Rabbit	Monkey	Human
Implantation	6	7	9	7
Primitive Streak	8–8.5	7	17–19	16–18
Neural Plate	9	8	19–20	18–20
First Somite	9–10	¾	20–21	20–21
First Branchial Arch	10	9.5	21–23	20
Ten Somites	10–11	9	23–24	25–26
Forelimb Buds	11	10–11	27–29	28
Hindlimb Buds	12	11	29–31	31–32
Forepaw Rays	14	15	35–37	37
Testis: Histological Differentiation	14–15	16–17	37–39	43–48
Heart: Septation Complete	16	17	36	46–47
Palate Closed	16–17	19–20	45–47	56–58
Gestation Length	22	32	165	267

*(Compiled from Shepard, 1989; Hoar and Monie, 1981; Nishimura and Shiota, 1977; Hendrickx et al., 1983).

Functional Maturation

Organogenesis is followed by growth and differentiation of function. Hence, Wilson (1973) hypothesized that the critical period for functional deficits would parallel that for functional maturation. However, subsequent studies on functional teratogenesis suggest a somewhat different pattern (Figure II-3) in which the peak sensitivity is still during organogenesis, but decreases only gradually through the final maturation of the organ. For example, administration of Cd to the rat during gestation days 12–15, the critical period for lung development, decreases lung weight but also impairs the maturation of the pulmonary surfactant system (Daston, 1982), an event that takes place on the last day of gestation. This functional effect is incompatible with postnatal life, as the newborns die of respiratory problems comparable to the human neonatal respiratory distress syndrome (Daston and Grabowski, 1979).

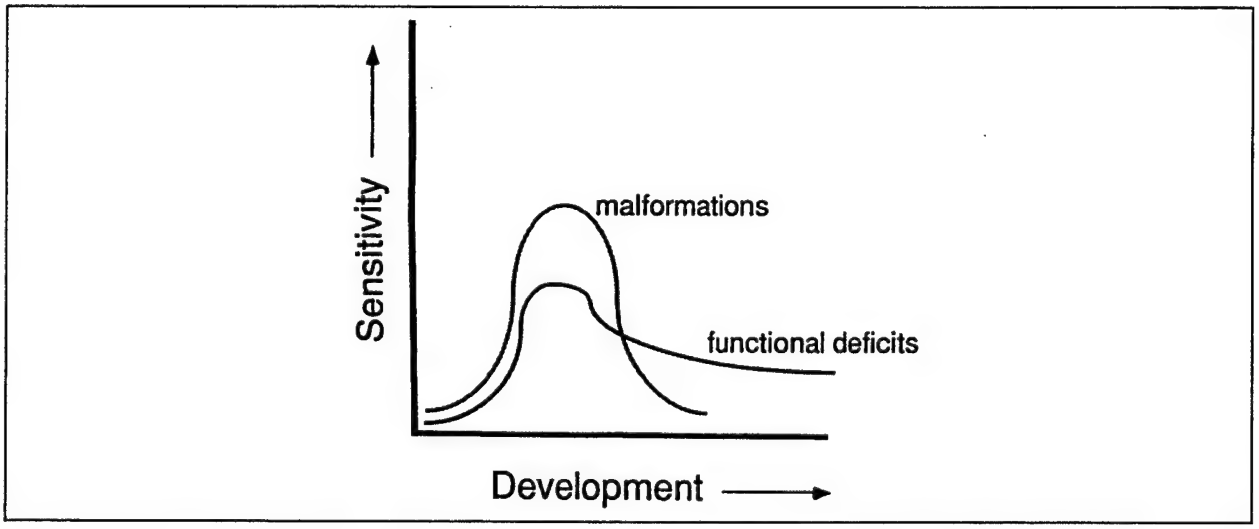


Figure II-3. The Critical Period for Functional Deficits begins at the Same Time as That for Structural Defects, but Extends through Maturation of the Organ. Sensitivity decreases with time.

Kidney development provides an illustration for the gradually decreasing sensitivity of an organ to agents causing functional deficit. Kidney development is characterized by a continuous differentiation of tubules—nephrogenesis—from the initial formation of the kidney to its ultimate maturation. Nephrogenesis extends well into the late fetal period in humans (Potter, 1972), and the second postnatal week in rats (Baxter and Yoffey, 1948). Intuitively, insult at earlier stages would cause a more severe effect, as more undifferentiated tubules would be present. Severity would gradually

diminish as the fraction of immature tubules decreased. This appears to be borne out by a review of the database on renal functional teratology (Lau and Kavlock, 1994), as the period of greatest sensitivity to persistent renal functional deficit coincides with that of kidney organogenesis. However, it is still possible to elicit persistent functional deficits with much later treatment. For example, colchicine administered to postnatal day 1 rats caused a persistent deficit in urine concentrating ability (Daston et al., 1988). [Interestingly, this long period of maturation makes the developing kidney less sensitive to adult renal toxicants, including mercuric chloride (Daston et al., 1983) and sodium fluoride (Daston et al., 1985). Unlike the functional teratogens, the sensitivity to these agents increases with maturity and seems to be a function of both maturity of the nephrons and increased distribution of toxicant to the kidney (Daston et al., 1986)]. This pattern of gradually decreasing sensitivity suggests that systems with prolonged maturation are more prone to functional deficit, a notion supported by the extreme sensitivity of the nervous system, which takes the longest time to mature.

Time of Exposure and Developmental Outcome: Figure II-4 presents three hypothetical dose–response relationships for the three possible adverse outcomes (embryonic death, malformation, growth retardation) evaluated in most developmental toxicity studies. Although the relationships among the three outcomes (i.e., proceeding from malformation and/or retardation to lethality) may be preserved across developmental stages, the proportions of individuals affected in each category, as well as the dosage needed to produce the outcome, are likely to change significantly and rapidly as development progresses. In other words, the critical period for structural defects ends with the end of organogenesis (with some exceptions), but growth is more sensitive during the fetal period, when the conceptus increases markedly in size. The dosage at which lethality occurs often varies with developmental stage as well, particularly if death is a consequence of a structural malformation whose development is restricted to a critical period (e.g., cardiac agenesis).

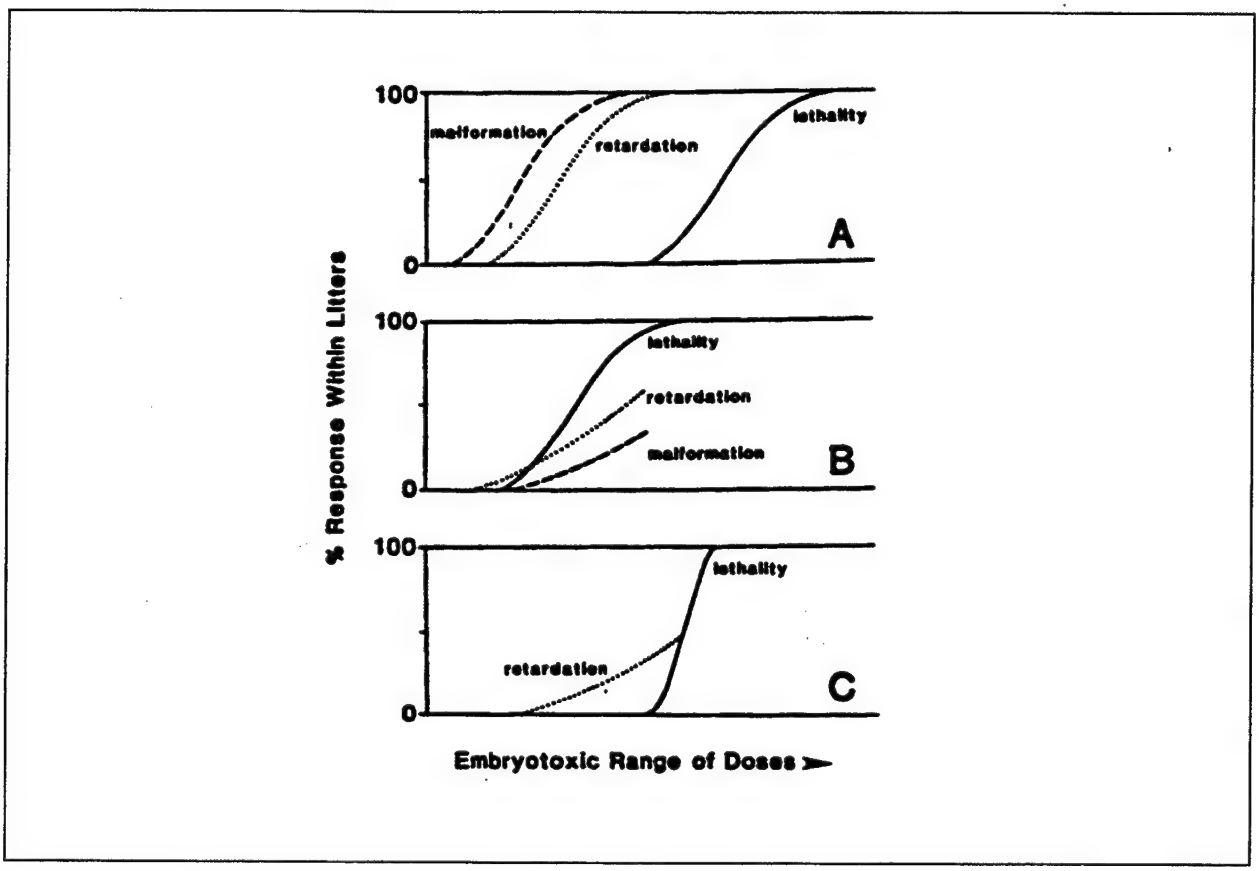


Figure II-4. Three Hypothetical Dose-Response Relationships for Malformation, Growth Retardation, and Death, the Three Developmental Manifestations Evaluated in Most Screening Studies. It is important to note that these responses are dependent not only on dosage but also on time of exposure. From Manson (1980), with permission.

Conclusions

The extremely rapid progression of the developing organism through a series of unique life stages accounts for a changing pattern of sensitivity to and outcome from toxic insult. Early stages are susceptible to embryolethality but largely resistant to teratogenesis; the organogenesis-stage embryo is extremely sensitive to the production of structural and, to some extent, functional abnormalities; and fetal and postnatal growth and maturation phases become increasingly refractory to structural, and eventually functional, alterations. These phases can be further subdivided into critical periods for the structural and functional development of each organ and organ system.

The uniqueness of development makes it an exceptional case for risk assessment. First, observed effects are only interpretable in the context of the time of exposure. This is important in determining the plausibility that a given exposure caused a certain effect. Second, as a single, ill-timed exposure is

sufficient to be adverse, risk assessment techniques such as exposure averaging, or reference doses for chronic exposures are not applicable to developmental effects.

REFERENCES

Barr, M. and Cohen, M.M. (1991). ACE inhibitor fetopathy and hypoclavaria: the kidney-skull connection. *Teratology* 44:485–495.

Baxter, J.S. and Yoffey, J.M. (1948). The postnatal development of renal tubules in the rat. *J. Anat.* 82:189–197.

Brent, R.L. and Beckman, D.A. (1991). Angiotensin-converting enzyme inhibitors, an embryopathic class of drugs with unique properties: information for clinical teratology counselors. *Teratology* 43:543–546.

Daston, G.P. (1982). Toxic effects of cadmium on the developing rat lung. II. Glycogen and phospholipid metabolism. *J. Toxicol. Environ. Health* 9:51–62.

Daston, G.P. (1994). Other functional abnormalities: methodology and data evaluation. *In: Handbook of Developmental Toxicology* (Hood, R.D., ed.), CRC Press, Boca Raton, FL (in press).

Daston, G.P. and Grabowski, C.T. (1979). Toxic effects of cadmium on the developing rat lung. I. Altered pulmonary surfactant and the induction of respiratory distress syndrome. *J. Toxicol. Environ. Health* 5:973–983.

Daston, G.P., Ebron, M.T., Carver, B., and Stefanadis, J.G. (1987). *In vitro* teratogenicity of ethylenethiourea in the rat. *Teratology* 35:239–245.

Daston, G.P., Kavlock, R.J., Rogers, E.H., and Carver, B. (1983). Toxicity of mercuric chloride to the developing rat kidney. I. Postnatal ontogeny of renal sensitivity. *Toxicol. Appl. Pharmacol.* 71:24–41.

Daston, G.P., Rehnberg, B.F., Carver, B., and Kavlock, R.J. (1985). Toxicity of sodium fluoride to the postnatally developing rat kidney. *Environ. Res.* 37:461–474.

Daston, G.P., Rehnberg, B.F., Carver, B., Rogers, E.H., and Kavlock, R.J. (1988). Functional teratogens of the rat kidney. I. Colchicine, dinoseb, and methyl salicylate. *Fundam. Appl. Toxicol.* 11:381–400.

Daston, G.P., Rehnberg, B.F., Hall, L.L., and Kavlock, R.J. (1986). Toxicity of mercuric chloride to the developing rat kidney. III. Distribution and elimination of mercury during postnatal maturation. *Toxicol. Appl. Pharmacol.* 85:39–48.

Generoso, W.M., Rutledge, J.C., Cain, K.T., Hughes, L.A., and Braden, P.W. (1987). Exposure of female mice to ethylene oxide within hours after mating leads to fetal malformation and death. *Mutat. Res.* 176:269–274.

Generoso, W.M., Rutledge, J.C., Cain, K.T., Hughes, L.A., and Downing, D.J. (1988). Mutagen-induced fetal anomalies and death following treatment of females within hours of mating. *Mutat. Res.* 199:175–181.

Hendrickx, A.G., Binkerd, P.E., and Rowland, J.M. (1983). Developmental toxicity and non-human primates: interspecies comparisons. *Iss. Rev. Teratol.* 1:149–180.

Hoar, R.M. and Monie, I.W. (1981). Comparative development of specific organ systems. In: *Developmental Toxicology* (Kimmel, C.A. and Buelke-Sam, J., eds.), Raven Press, New York, pp. 13–33.

Lau, C. and Kavlock, R.J. (1994). Renal and other organ system dysfunction. In: *Developmental Toxicology* (Kimmel, C.A., and Buelke-Sam, J., eds.), Raven Press, New York, pp. 119–188.

Manson, J.M. (1980). Teratogens. In: *Casarett and Doull's Toxicology: The Basic Science of Poisons, 3rd ed.* (Klaassen, C.D., Amdur, M.O., and Doull, J., eds.), Macmillan Publishing, New York, pp. 195–220.

Markert, C.L. and Petters, R.M. (1978). Manufactured hexaparental mice show that adults are derived from three embryonic cells. *Science* 202:56–58.

Martin, R.A., Jones, K.L., Mendoza, A., Barr, M., and Benirschke, K. (1992). Effect of ACE inhibition on the fetal kidney: decreased renal blood flow. *Teratology* 46:317–321.

Nishimura, H. and Shiota, K. (1977). Summary of comparative embryology and teratology. In: *Handbook of Teratology*, vol. 3 (Wilson J.G. and Fraser, F.C., eds.), Plenum Press, New York, pp. 119–154.

Potter, E.L. (1972). Normal and Abnormal Development of the Kidney, *Year Book Medical Publishers, Chicago.*

Rogers, J.M., Chernoff, N., and Shuey, D.L. (1994). Postclosure neural tube defects induced by ethylenethiourea (ETU) in the rat. *Teratology* 49:383.

Ruddick, J.A. and Khera, K.S. (1975). Pattern of anomalies following single oral doses of ethylenethiourea to pregnant rats. *Teratology* 12:277–282.

Rutledge, J.C. and Generoso, W.M. (1989). Fetal pathology produced by ethylene oxide treatment of the murine zygote. *Teratology* 39:563–572.

Shepard, T.H. (1989). *Catalog of Teratogenic Agents,* 6th ed., Johns Hopkins Univ. Press, Baltimore.

Wilcox, A.J., Weinberg, C.R., Wehmann, R.E., Armstrong, E.G., Canfield, R.E., and Nisula, B.C. (1985). Measuring early pregnancy loss: laboratory and field methods. *Fertil. Steril.* 44:366–374.

Wilson, J.G. (1973). *Environment and Birth Defects.* Academic Press, New York.

EXPOSURE-DURATION RELATIONSHIPS: THE RISK ASSESSMENT PROCESS FOR HEALTH EFFECTS OTHER THAN CANCER

Gary L. Kimmel, Ph.D.
Reproductive and Developmental Toxicology Branch
Office of Research and Development, U.S. Environmental Protection Agency

ABSTRACT

The development of the risk assessment process over the past ten years has not only provided us with a framework to carry out an important regulatory function, but has aided in the identification of major gaps in our knowledge and the scientific data base. One of those gaps is the influence of the duration of exposure on the toxicity of an agent. Currently, approaches to estimating short-duration exposure limits are only generally defined, assuming a constant relationship between the level of an exposure and its duration with respect to the response that can be expected. Within developmental toxicology, even this type of exposure duration adjustment is generally not carried out, the risk assessment being based on the overall daily exposure, regardless of the actual timing, duration or frequency of exposure or the physical-chemical properties of the agent.

This paper will summarize several current approaches to estimating short-duration exposure limits and outline preliminary model development that may be useful in better defining the influence of exposure duration on toxicity. The focus is on developmental toxicity, but most of the concepts should be generally applicable to other non-cancer health endpoints.

INTRODUCTION

When the National Research Council outlined the risk assessment process (National Research Council, 1983), it helped put into perspective the relationship of toxicity testing, data analysis, risk characterization, and the risk management process. This effort was directed primarily at cancer, but much of the framework is applicable to other endpoints of toxicity as well. As a result, the U.S. Environmental Protection Agency (EPA) developed specific risk assessment guidelines that not only identified methodology that could be applied to the assessment of toxicity data, but also identified gaps in our knowledge (U.S. EPA, 1986a; 1991). While approaches for incorporating less-than-lifetime

exposures in the regulatory process have been developed, our understanding of the influence of timing and duration of exposure on the toxicity of chemicals is limited at best.

The *Conference on Temporal Aspects of Risk Assessment for Noncancer Endpoints* provided an important forum for scientists, risk assessors, and risk managers to examine exposure-related issues in the context of specific areas of chemical toxicity or exposure scenarios. One such focal area was developmental toxicity. The importance of critical periods in development and the changing susceptibility of the embryo/fetus to exposure were reviewed in the previous paper by Daston and Manson. This paper addresses another temporal aspect, namely, exposure duration. Several current approaches to estimating short-duration exposure limits are summarized, and preliminary model development that may be useful in better defining the influence of exposure duration on toxicity is outlined. The examples are drawn largely from the author's experience within the EPA.

CURRENT APPROACHES TO INCORPORATING DURATION OF EXPOSURE IN STANDARD SETTING

In the past, most regulatory agencies have placed primary emphasis on the potential health effects from lifetime exposures. However, there is an increasing recognition among regulatory agencies that continuous lifetime exposures are more often the exception, than the rule. Intermittent occupational, recreational or accidental exposures are much more likely.

Table 1 notes just some of the EPA regulatory program office and regional concerns that have been raised with regard to acute or short-term exposures. These concerns span a number of program offices, as well as applications from advisory to standard setting. The Office of Drinking Water develops health advisories for 1-day and 10-day consumption levels, considering exposures to both adults and children. Many of the inhalation exposures that the Office of Air and Radiation is concerned with in establishing standards are associated with occupational or recreational exposures that are limited in both duration and frequency. Other concerns relate to short- or intermediate-term clean ups and emergency releases, and episodic exposures to agents such as pesticides, volatile organics, etc. In each of these examples, an assumption of a lifetime exposure scenario is inconsistent with the known or expected human exposure and may be meaningless in establishing health-based standards.

Table II-3. U.S. EPA Regulatory Concerns over Short-Duration Exposures.

Office of Water
Drinking Water Health Advisories
Office of Air and Radiation
Inhalation Standards
Office of Solid Waste and Emergency Response/Regions
Hazardous site Clean Up
Office of Prevention, Pesticides and Toxic Substances
Emergency Planning, SARA Title III
Pesticide Application

The importance of incorporating temporal aspects into noncancer risk assessment has been recognized for years, but the process to accomplish this has, for the most part, been only generally defined. Table II-4 lists several examples of approaches that have been developed for establishing short-duration exposure limits. The threshold limit value for short-term exposure limits (TLV-STEL) is one of the more recognized of these approaches. Established by the American Conference of Governmental Industrial Hygienists, the TLV-STEL is the time-weighted average concentration to which a worker can be exposed for fifteen min in an 8-h workday without experiencing irritation, irreversible tissue damage or impaired work (ACGIH, 1986). Apparently this concentration can be reached as many as 4 times a day, if periods of 60 min occur between each time (Ayers and Taylor, 1989). The threshold limit value for excursion limits (TLV-EL) is calculated when insufficient data exist to determine a TLV-STEL. The TLV-EL equals a concentration that is 3 times the standard TLV, is based on a time-weighted average, and should not be exceeded for more than 30 min during the workday. At no time should the TLV be exceeded by fivefold. The TLV-STEL focuses on protecting workers from brief, but potentially recurring, workplace exposures, and the endpoints that are considered include irritation and impairment of work. Other approaches are designed for different exposure scenarios, and consequently, focus on other endpoints and expected exposure durations. For example, Immediately Dangerous to Life and Health (IDLH) values developed by the National Institute for Occupational Safety and Health are intended to be protect workers against accidental exposures (30 min or less) that may be lethal or cause irreversible health effects. The Emergency Exposure Guidance Levels (EEGLs) established for emergency exposures of military personnel where moderate impairment of performance that does not affect an emergency response is permissible (Committee on Toxicology, 1986). The duration of the expected exposure can be as much as 24 h, but EEGLs are

developed for different exposure durations, including less than an hour. When data is limited for developing these values, "Haber's law," which assumes a constant product of concentration times duration, is applied. Thus, exposure to a unit of concentration for 10 h is considered identical to an exposure to a 10 unit concentration for 1 h.

Table II-4. Guidelines for Estimating Short-Duration Exposure Limits.

	Endpoints	Duration
TLV-STEL	Irritation, Irrev. Tissue Damage, Impaired Work	0.25 hrs
TLV-EL	≤ 3× TLV – TWA	0.5 hrs
IDLH	Death, Irrev. Health Effects	
EEGL	Acutely Debilitating or I H I — Military	<1–24 hrs
CEEL	Adverse Health Effects	1–8 hrs
SPEGL	Adverse Health Effects — Public	1–24 hrs
HA	Adverse Health Effects	1–10 days

TLV-STEL: Threshold Limit Value-Short-Term Exposure Limit.
TLV-EL: Threshold Limit Value-Excursion Limits.
IDLH: Immediately Dangerous to Life and Health.
EEGL: Emergency Exposure Guidance Levels
CEEL: Community Emergency Exposure Levels.
SPEGL: Short-Term Public Emergency Guidance Levels.
HA: Health Advisory

The other guidelines listed in Table II-4 are applicable to the general public and therefore, all adverse health effects are considered in setting a standard. The Community Emergency Exposure Levels are maximum inhalation exposure values for the public, "applicable to emergency exposures of foreseeable magnitude and duration (usually 1 to 8 hours, although few exposures will exceed 1 hour)." (Committee on Toxicology, 1993). Short-Term Public Emergency Guidance Levels are used by the military for emergency exposures, similar to EEGLs, but for the general public. Finally, as noted above, the EPA's Health Advisories are developed by the Office of Drinking Water to provide technical guidance on drinking water contamination levels that may result in health effects. These advisories are developed for exposure periods of 1-day and 10-days, considering consumption characteristics applicable to adults as well as those applicable to children.

In many ways, developmental toxicity presents a somewhat unique example of exposure duration issues as they relate to health endpoints other than cancer. In the previous paper, Daston and Manson reviewed the important influence that the stage of development may have on whether an exposure will

lead to an effect. In the developing animal, cells, tissues, and organ systems are part of an ever-changing environment, and if an exposure of sufficient magnitude occurs during a critical period in the development of a particular system, then an effect is likely to be observed. However, if the development of the organ system is not at a stage that can be affected, then an alteration of development will not occur, regardless of the exposure. The nature of the critical period to a large extent defines the exposure conditions necessary to result in altered development, including the exposure concentration, duration, frequency, etc. In reality, this is not unique to developmental toxicology, for adult systems should also be viewed in the context of their changing responsiveness with time and under different conditions. Consequently, this and other concepts employed in developmental toxicology should be useful in other fields of toxicology and vice versa.

Developmental toxicity risk assessment, as described in the EPA Guidelines (U.S. EPA, 1991), provides an example of how the characteristics of exposure duration are considered in the assessment process. Typically, if the database is adequate, a reference dose or reference concentration for developmental toxicity (RfD_{DT}; RfC_{DT}) is derived in a manner similar to RfDs and RfCs for chronic exposure, using the identified no-observed-adverse-effect-level and uncertainty/modifying factors (Kimmel, C. and Kimmel, G., 1994). A major exception to the chronic RfD/RfC is that the RfD_{DT}/RfC_{DT} recognizes that developmental effects may result from single exposures and that long-term exposure is not factored into the derivation (U.S. EPA, 1986b; 1991). The statement of this assumption in the original guidelines, began to move our thinking away from the traditional long-term or lifetime, low-dose exposure scenario that still exerts an undue influence on the regulatory process.

Even with this recognition of the importance of evaluating less-than-lifetime exposures, little has been done in any risk assessment or regulatory guidance to specifically incorporate the influence of exposure duration into the methods of risk evaluation. Within developmental toxicity risk assessment, the exposure basis for calculation is a daily exposure, regardless of the actual timing, duration or frequency of exposure or the physical-chemical properties of the agent (Kimmel, C. and Kimmel, G. 1994). A 1-h inhalation exposure is ultimately viewed the same as a 6-h inhalation exposure, and both are considered similar to a bolus gavage study, where the agent must pass through the gastrointestinal barrier and hepatic circulation before moving into the systemic circulation. Seldom is cumulative exposure over the experimental period factored in, even though the standard developmental toxicity test uses an exposure period that covers all of mid-gestation (e.g., gestation Days 6–15 in the rat). Of course, the risk assessment guidance is to exercise scientific judgment in all aspects of the process,

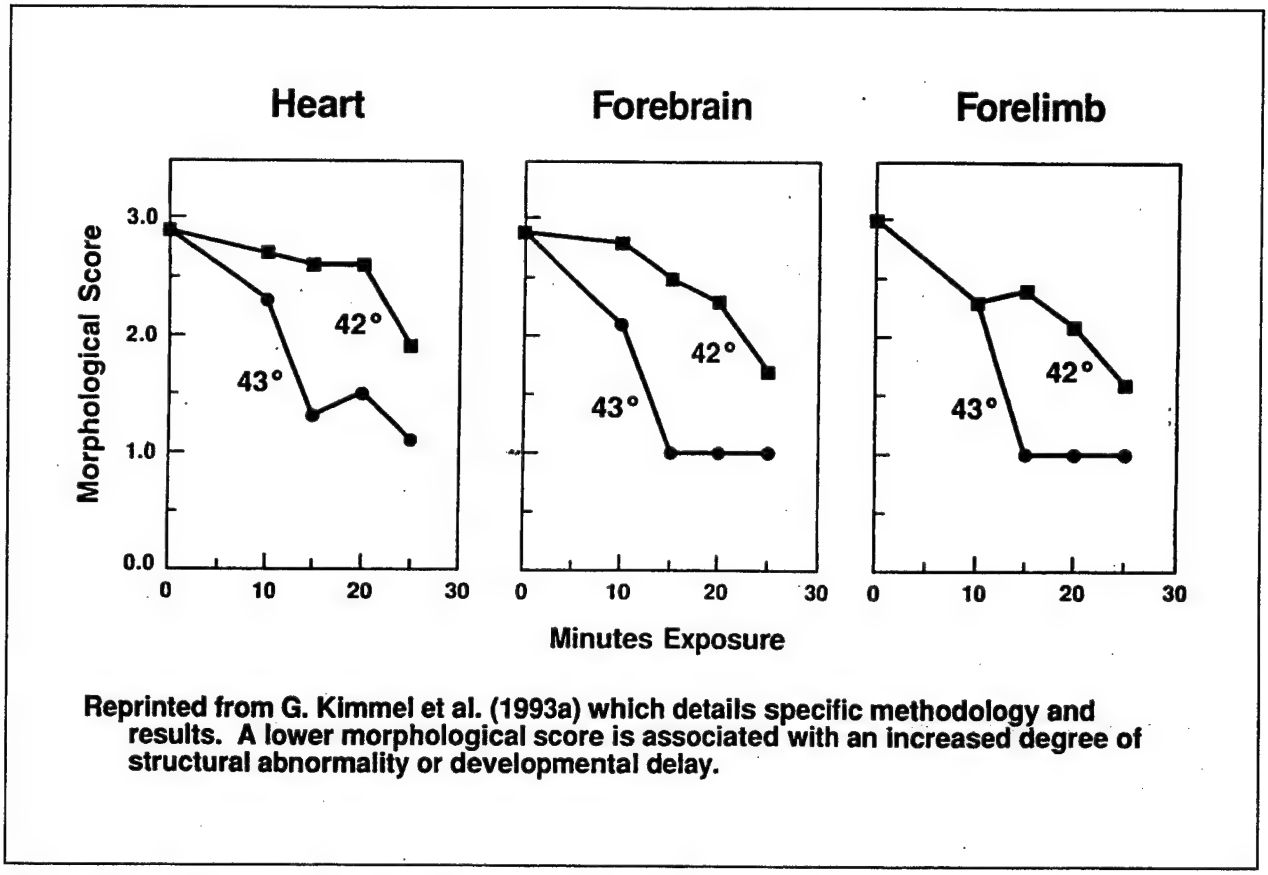
including defining the exposure conditions and their impact on the data (U.S. EPA, 1991). The developmental toxicity risk assessment guidelines encourage the use of pharmacokinetic data where possible, recognizing that it is rarely available.

MODEL DEVELOPMENT

In response to these limitations of the current procedures, the Office of Research and Development, through its program on *Research In Health Risk Assessment*, is beginning to focus more attention on developing biologically based dose–response models for risk assessment (U.S. EPA, 1990). One specific area of interest is improving "our understanding of the interplay among rate, intensity, and duration of exposure as they affect toxicological outcomes...." In conjunction with this program and the Office of Air and Radiation, our laboratory has initiated studies on the relationship of exposure duration and developmental effects.

Using whole embryo cultures, studies on hyperthermia and its effect on intrauterine development in the rat demonstrated that the effect was dependent both on the exposure temperature and the duration (Figure II-5) (Kimmel, G. et al., 1993a). The shape of the response curves suggested a potential for mathematical modeling that might be applied to testing the concept of "Haber's Law" (i.e., constant concentration x duration combinations yield similar responses). Testing this concept is difficult in developmental toxicology, in large part because of the constantly changing biological entity under study, as noted above, but also because of the difficulty in approaching steady state conditions for the purpose of comparisons (Smith and Clark, 1987).

The use of hyperthermia as a model exposure provides distinct advantages for addressing this concept and previous studies have indicated a close relationship between the findings *in vivo* and *in vitro* (Kimmel, C. et al., 1993; Kimmel, G. et al., 1993a; Cuff et al., 1993) In whole embryo culture, where the embryo is isolated from the maternal animal and maintained in culture over 24 h, specific temperature levels can be rapidly achieved and maintained over extended durations. Hyperthermia is a physical agent and is not influenced by or dependent on the embryo's pharmacokinetic characteristics. Possibly of greatest importance is the fact that hyperthermia causes observable effects on *in vitro* development after only brief (minutes) exposures, and there is an increase in these effects as the temperature and/or duration are increased over the first 60 min of exposure. The normal progression in development that takes place during this 60-min period can be considered minimal, thus permitting a direct comparison of the concentration x duration concept on well-defined endpoints.



Reprinted from G. Kimmel et al. (1993a) which details specific methodology and results. A lower morphological score is associated with an increased degree of structural abnormality or developmental delay.

Figure II-5. Examples of the Response Duration Pattern for the Heart Forebrain and Forelimb following Exposure to 42 or 43 °C.

Extension of the original *in vitro* hyperthermia findings has employed a response-surface modeling technique (Box and Draper, 1987). A response-surface model is a mathematical technique for assessing biological responses with a combination of two or more variables, using a "response-surface" (Figure II-6). In the situation where the response is dependent solely on a constant temperature x duration combination, the surface would appear as it does in Figure II-6. The response would increase symmetrically from control values as either temperature or duration were increased. However, if there was an interaction between elevations in temperature and duration of exposure, then the response-surface would not demonstrate bilateral symmetry and additional parameters would have to be considered in the model.

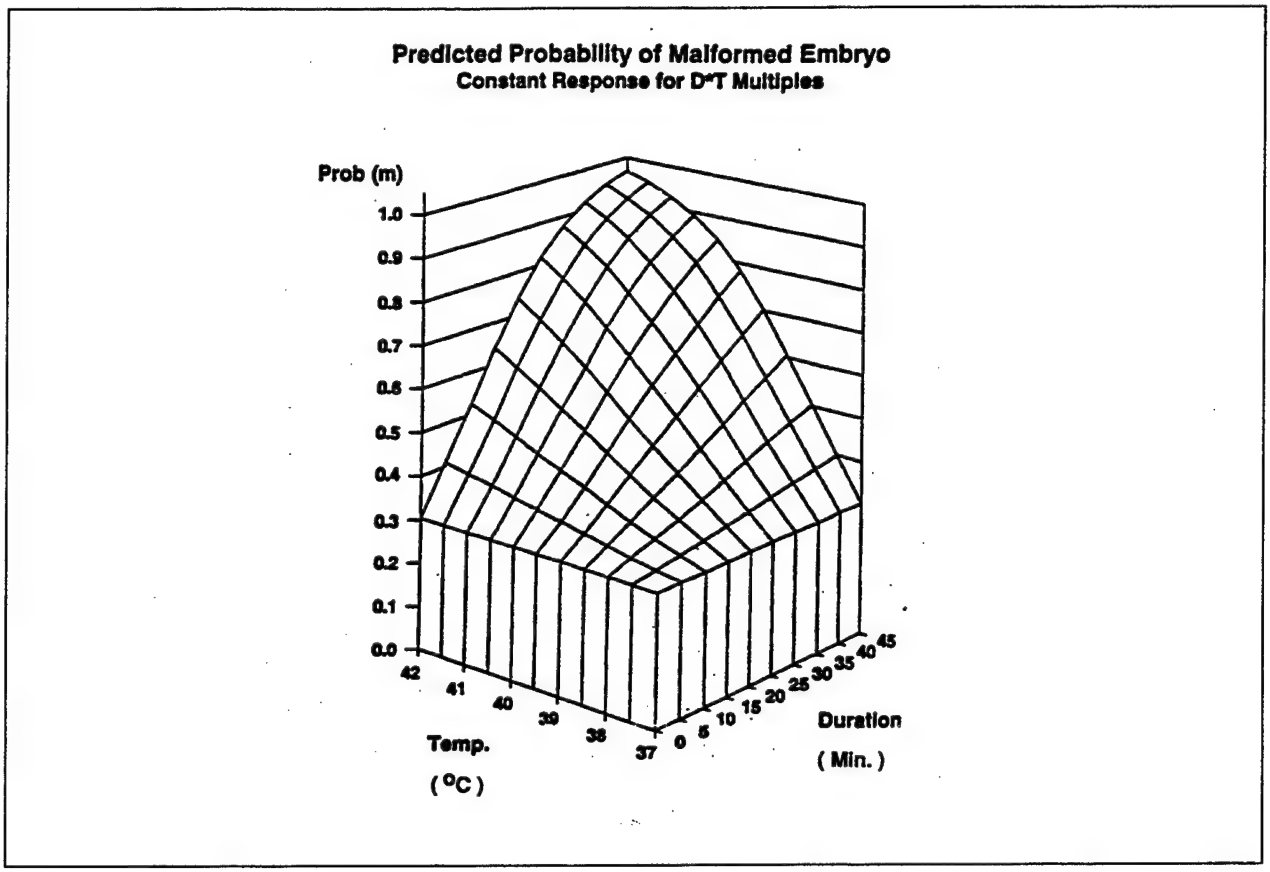


Figure II-6. Response Surface Methods. Response surface methods provide graphical and analytical tools for assessing the *combined effect* of more than one variable on a particular response.

The preliminary findings (Kimmel, G. et al., 1993b) indicate that for endpoints currently being modeled (i.e., embryo viability, developmental abnormality, and total number of embryos affected), a constant temperature x duration combination does not define the responses of the embryo, indicating that "Haber's Law" does not hold in this situation. However, the database needs to be expanded considerably. One of the long-range goals of this effort is to evaluate the appropriate experimental design for studying exposure-duration effects. Even the limited database that has been developed indicates that developing the initial model and describing the response-surface requires a considerable number of data points throughout the range of temperatures and durations being studied. This data requirement will undoubtedly be reduced as our understanding of the parameters and their interrelationship increases. However, the applicability of such approaches in the risk assessment process will in part be based on their utility with limited databases.

SUMMARY

The regulatory community's experience with risk assessment, especially as it relates to noncancer health effects, has clearly pointed out a need to examine the potential toxicity of exposures that are less-than-lifetime and potentially singular or episodic in nature. There is a growing understanding that the manner in which these types of exposures are considered in setting exposure limits is simplistic; that as biologically based dose–response models are developed, exposure characteristics such as timing, duration, and frequency must be factored into the risk characterization. This paper has attempted to bring some of this into perspective with regard to exposure duration, and to provide one example of the type of modeling efforts that may be possible to address some of these issues.

ACKNOWLEDGMENTS

The author wishes to recognize the important contributions of the other principal investigators in the response-surface modeling studies, including those of Dr. P.L. Williams, Dr. C.A. Kimmel, Dr. L.M. Ryan, and Mr. T.W. Claggett. The author also thanks Drs. B.R. Sonawane and T.M. Crisp for their helpful suggestions during the preparation of this manuscript.

REFERENCES

ACGIH. 1986. Threshold limit values and biological exposure indices. *American Conference of Governmental Industrial Hygienists*, 5th edition, Cincinnati, Ohio.

Ayers, P.H. and Taylor, W.D. 1989. Solvents. In: *Principles and Methods of Toxicology*, ed. A.W. Hayes. New York: Raven Press.

Box, G. and Draper, N. 1987. *Empirical Model Building and Response Surfaces*. New York: John Wiley and Sons.

Committee on Toxicology. 1986. *Criteria and methods for preparing emergency exposure guidance levels (EEGL), short-term public emergency guidance level (SPEGL), and continuous exposure guidance level (CEGL) documents.* National Academcy Press, Washington, DC.

Committee on Toxicology. 1993. *Guidelines for developing community emergency exposure levels for hazardous substances.* National Academcy Press, Washington, DC.

Cuff, J.M., Kimmel, G.L., Kimmel, C.A., Heredia, D.J., Tudor, N., and Chen, J. 1993. Relationship between abnormal somite development and axial skeletal defects in rats following heat exposure. *Teratology* 48:259–266.

Kimmel, C.A., Cuff, J.M., Kimmel, G.L., Heredia, D.J., Tudor, N., Silverman, P.M., and Chen, J. 1993. Skeletal development following heat exposure in the rat. *Teratology* 47:229–242.

Kimmel, C.A., and Kimmel, G.L. 1994. Risk assessment for developmental toxicity. In *Developmental Toxicology: Target Organ Toxicity Series*, eds, C.A. Kimmel and J. Buelke-Sam, 2nd edition. New York: Raven Press.

Kimmel, G.L., Cuff, J.M., Kimmel, C.A., Heredia, D.J., Tudor, N., and Silverman, P.M. 1993a. Embryonic development *in vitro* following short-duration exposure to heat. *Teratology* 47:243–251.

Kimmel, G.L., Williams, P.L., Ryan, L.M., Kimmel, C.A., and Tudor, N. 1993b. The effects of temperature and duration of exposure on *in vitro* development and response-surface modeling of their interaction. *Teratology* 47:401 (Abstract 64).

National Research Council. 1983. Risk assessment in the federal government: Managing the process. National Academy Press, Washington, DC.

Smith, D.A., and Clark, B. 1987. Pharmacokinetics and toxicity testing: Basic principles and pitfalls. In: *Pharmacokinetics in Teratogenesis,* Vol. I, eds, H. Nau and W.J. Scott, Jr. Boca Raton: CRC Press.

U.S. EPA. 1986a. The risk assessment guidelines of 1986. EPA/600/8-87/045, Washington, DC.

U.S. EPA. 1986b. Guidelines for the health assessment of suspect developmental toxicants. *Federal Register* 51: 34028–34040.

U.S. EPA. 1990. Research to improve health risk assessments (RIHRA) program. EPA/600/9-90/038, Washington, DC.

U.S. EPA. 1991. Guidelines for developmental toxicity risk assessment. *Federal Register* 56: 63798–63826.

REVERSIBILITY OF EFFECTS: OVERVIEW AND REPRODUCTIVE SYSTEMS[1]

Eric D. Clegg, Ph.D.
Reproductive and Developmental Toxicology Branch, Human Health Assessment Group
U.S. Environmental Protection Agency

SUMMARY

Recovery from adverse non-cancer health effects may occur in some circumstances after cessation of exposure or after adaptation to continuing low dose exposure. In this overview, factors that influence the reversibility of toxic effects of environmental chemicals are presented, using examples derived from toxicity to the reproductive system. Aspects that must be considered include exposure scenario, stage of the life cycle at which exposure occurs, and the nature of the toxicity. Selected categories of reproductive effects are considered briefly with respect to potential for recovery. Finally, the incorporation of reversibility into risk assessments is discussed.

INTRODUCTION

Traditionally, testing protocols and risk assessments for environmental agents have focused on effects of long-term exposure. With continuing exposure, recovery from an adverse non-cancer health effect is unlikely and the occurrence of an adverse effect at any time is cause for concern. Little attention has been directed to reversibility of effects when cessation of exposure makes recovery possible. That situation is changing with increased interest in risk assessments based on shorter-duration exposures. A scenario that allows recovery raises the issue of whether reversible effects warrant lesser concern than ones from which full recovery is not expected. If reduced concern is possible, then it is necessary to explore how a lower level of concern can be expressed in quantitative risk assessment and in risk management.

The intent of this paper is to provide an overview of the topic and "set the table" for further debate and action. It discusses the factors that should be considered in assessing a potentially reversible effect, as well as how reduced concern might be incorporated into a reference dose (RfD) or reference concentration (RfC) determination. Reproductive system examples are used, and the potential for

[1] Disclaimer: The views expressed in this paper are those of the author and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency. The U.S. Government has the right to retain a nonexclusive royalty-free license in and to any copyright covering this paper.

recovery from selected toxic effects on the female and male reproductive systems is discussed briefly. A book has been published previously that provides background information on reversibility of effects on the testis (Scialli & Clegg, 1992). Also, female and male reproductive toxicology have been discussed extensively in three recent publications (Manson & Kang, 1994; Zenick et al., 1994; Kimmel, G.L. et al., 1995).

REVERSIBILITY AND RECOVERY

In toxicology, reversibility refers to potential for exposed organisms to recover from the effects of toxic exposures to a state that would have existed without a toxic insult. Recovery may be attained either after cessation of exposure or by adaptation via a compensatory mechanism with continuing exposure to low doses. For example, recovery of normal ovulation would be expected in the next cycle following cessation of a short-term exposure to chlordimeform that interfered with the ovulatory surge of luteinizing hormone (Goldman et al., 1991; Stoker et al., 1991). Adaptation in the presence of continuing exposure is considered usually to result from induction of enzymes such as the cytochrome P-450 isozymes or UDP-glucuronosyl transferases that increase ability of the animals to metabolize the toxic compound. Whether adaptation should be considered equivalent to recovery is a debatable issue that may depend on the type of effect.

In the presence of apparent recovery from the original effect, a residual effect could remain that was not detectable by the array of end points being used. Fertility could be restored in female rodents even though a significant reduction in number of ovarian primordial follicles had occurred. Unless quantitative measurements of those follicles were done, that residual effect would go undetected. This example with ovarian follicles is pertinent for residual effects since those follicles cannot be replaced, and depletion of that cell population results in premature menopause in women (see Kimmel, G.L. et al., 1995). If undetected at the time(s) of exposure, the loss of follicles could be cumulative with each toxic exposure and not suspected until premature menopause occurred. It is possible also for reduced resistance to a subsequent insult to remain as a residual effect.

FACTORS AFFECTING REVERSIBILITY

Major factors that influence whether an effect is reversible are listed in Table II-5. These components are strongly interdependent and should not be considered in isolation.

Exposure

The potential that an adverse effect could be reversed is dependent, in part, on the conditions of exposure: how much, how long, and how frequently. The severity of a given effect is likely to increase with increasing dose level as is the probability that other effects may appear. With increasing severity of an effect, the probability also increases that the effect will be irreversible. Thus, numerous testicular toxicants at low dose levels affect only the spermatocyte population initially, but at high doses can affect the spermatogenic stem cell population and its ability to repopulate the seminiferous epithelium. Spermatocytes can be replaced, but significant damage to the stem cell population is unlikely to be reversible (see Nature of Toxicity).

The duration of exposure may range from very brief to near lifetime. If exposure ceases, the time required for clearance of the toxic compound or a toxic metabolite from the target organ and repair of the effect become important. Time for clearance of a compound is a function of the level in the tissue, availability of a continuing source via mobilization from another tissue (e.g., adipose tissue for lipophilic compounds), and capability to remove the compound via metabolism, circulation and excretion. Dose level, type and severity of effect induced, time to clear the toxicant, and time to repair the effect combine to determine the capability and timing of recovery.

Exposure scenarios that involve intermittent exposures present challenges that may not be practical to address given the current level of information. Assessments of such situations should consider not only the duration and frequency of exposures, but also the duration of the interval(s) between exposures and the time required for recovery from the effect(s). Those factors, as well as the level of each exposure, could vary for each repetition of exposure and between exposed individuals. Further, the potential could exist for cumulative residual damage with increasing repetitions. Worse yet, some situations, such as with pesticide applicators, may involve sequential exposure to different toxicants. Taken together, the assumptions and compromises that would be necessary for an intermittent-exposure assessment would probably result in an unacceptable level of uncertainty except in unusually simple cases. Thus, it may be prudent to treat such complex exposure patterns as if they were continuous.

Table II-5. Factors Affecting Reversibility.

Exposure	Stage of Life Cycle	Nature of Toxicity
Level	Development	Cell Type Affected
Duration	Adult	Severity of Cellular Effect
Pattern		Ability of Cells to Recover
Accumulation		
Clearance		

Period of life cycle

The period of the life cycle at which exposure occurs can affect both the type of effect that is induced and the degree to which a given effect is expressed. Here, attention is directed particularly at reproductive effects induced during development in contrast to those caused in postpubertal adults.

In the fetus, the processes of cell differentiation, migration and organ development are particularly susceptible to disruption. Interference with those processes results generally in deficiencies that are considered irreversible (see Development). Higher doses of the same agent may be required to produce an adverse effect in adults, and the adult effect might be reversible.

Estrogenic chemicals provide a good example. When pregnant females are exposed to sufficient exogenous estrogen, male offspring have higher incidences of undescended testes, reduced sperm production, hypospadia and reproductive system cancer (reviewed in Sharpe & Skakkebaek, 1993). The reduced sperm production and cancer are not detected until later in life. Reduced sperm production resulting from developmental exposure of males to excess estrogen is due to interference with proliferation of Sertoli cells that occurs only in the fetal and neonatal testes. Since Sertoli cells are required to support spermatogenesis postpubertally, the number of sperm that can complete spermatogenesis is reduced. It is not clear whether the number of gonocytes (precursers to spermatogenic stem cells) may also be reduced, but the histologic appearance of the postpubertal testes is normal qualitatively. When adult males are treated with higher levels of estrogens, spermatogenesis may be impaired, but moderate impairment is likely to be reversible.

It is fascinating that when pregnant rats are treated with propylthiouracil to produce a hypothyroid condition, Sertoli cell proliferation in the male fetuses proceeds beyond the normal level (Cooke et al., 1992). The male offspring resulting from such treatment have enlarged, but morphologically normal, testes that produce more sperm than controls.

Nature of toxicity

When the concentration of toxicant at the target site has been reduced below an effective level, recovery is dependent on the cell type(s) affected, the severity of sublethal cellular toxicity, the ability of those cells to repair the damage, and, if lethal to cells, the ability to regenerate those cells. These factors will be illustrated by considering the disruption and recovery of spermatogenesis by two testicular toxicants.

The toxic effects of ethylene glycol monomethyl ether (EGME) and boric acid (BA) on spermatogenesis have been studied extensively. These chemicals are also members of only a small group of environmental chemicals for which recovery from reproductive effects has been examined. Even with these chemicals, key information is missing that is needed to judge the level of damage from which complete recovery is possible.

EGME appears to be primarily a germ cell toxicant (Foster et al., 1983; Chapin et al., 1985), while BA (Chapin & Ku, 1994) appears to affect Sertoli cells, as well as developing germ cells toxicant. However, because of the role of Sertoli cells in supporting spermatogenesis, toxic effects on Sertoli cells are likely also to interfere secondarily with that process. Thus, while both chemicals can disrupt spermatogenesis, different cell types are affected initially and the consequences for recovery from adverse effects may be different.

As would be expected, the severity of effect increases with higher doses. When male rats were exposed to a low dose level of EGME, the first manifestation of toxicity was damage to the spermatocytes. At higher doses, the "window of effect" became wider so that differentiating and proliferative spermatogonia (and possibly stem cell spermatogonia) as well as round spermatids were also affected directly (Chapin et al., 1985). In contrast, Sertoli cells are involved in the toxicity of BA. Effects of BA on Sertoli cells in rats probably cause the failure of sperm release from the Sertoli cells (spermiation) that is the first effect seen at the light microscope level. That effect is followed by destruction of spermatocytes and spermatids. An important distinction is that spermatocytes or spermatids are replaceable cells whereas Sertoli cells are not. Therefore, if Sertoli cells are damaged to the point where they cannot regain normal function, spermatogenic capability will be impaired irreversibly. Damage to any spermatogenic cell type beyond the stem cell spermatogonia should be reparable, but if stem cell spermatogonia are completely destroyed, the effect is not reversible. It is unclear to what degree the stem cell population can be affected and still retain the ability to recover fully.

When effects are limited to loss of differentiated spermatogenic cells and/or sublethal toxicity to Sertoli cells, it has been assumed generally that full recovery was likely. However, the results of recovery studies with BA (Chapin & Ku, 1994) do not support that stance. With BA, spermatogenesis did not resume where there was significant tubular atrophy despite the presence of spermatogonia. Those results suggest that the ability of Sertoli cells to support spermatogenesis was not restored fully. Further, it is instructive that quantification of the effect on number of remaining cells was necessary to detect the failure of spermatogenesis to recover fully. In seminiferous tubules that had not been affected to the point of severe atrophy, the histopathologic appearance after the recovery period was indistinguishable qualitatively from controls.

The timing for recovery of spermatogenesis involves several factors. First, consideration must be given to the spermatogenic cell types affected and the severity of that effect. The earlier in the spermatogenic process that an effect occurs, the longer it takes for the adverse effect to appear in ejaculates and the longer it takes subsequently for normal sperm to reappear (see Russell et al., 1990).

Repair of more severe effects, if repair is possible, also takes longer. With increasing dose level of BA, the observed effect progressed from impairment of spermiation to destruction of spermatids and spermatocytes. Any recovery seen in that latter situation would take longer than recovery from impairment of spermiation because increased time for restoration of the seminiferous epithelium once the toxicant was removed. With some chemicals, it also could take longer to reduce the level of toxicant in the testes. In this example with boric acid, qualitative restoration of the seminiferous epithelium was dependent on repair of the sublethal damage to the Sertoli cells followed by regeneration of the populations of spermatocytes and spermatids.

REVERSIBILITY OF REPRODUCTIVE EFFECTS

The reproductive systems of non-pregnant females, pregnant females and males involve complex interactions of the reproductive organs with the hypothalamus and pituitary, and, to a lesser extent, with other endocrine and non-endocrine organs. Table II-6 lists major components of the reproductive process that may be assessed. In testing for reproductive toxicity with laboratory species, each component is included in testing by using end points that reflect the culmination of those complex interactions. Those end points are presented in the EPA Draft Guidelines for Reproductive Toxicity Risk Assessment (U.S. Environmental Protection Agency, 1994) as well as by Zenick et al., (1994) and Kimmel et al. (1995). Selected elements of those components are discussed below with respect to

potential for recovery from adverse effects. The complexity of the subject dictates that this examination be superficial.

Development

The EPA Guidelines for Developmental Toxicity Risk Assessment (U.S. Environmental Protection Agency, 1991) indicate four categories of developmental effects that can occur. These are fetal death, malformation, impaired growth or behavioral deficits. Obviously, fetal death or malformation are irreversible effects. Under some conditions, impaired growth or a behavioral deficit could be repaired. However, in the absence of specific information that is not available usually, it is not considered possible to determine which of the manifestations of developmental toxicity that might be seen in testing with laboratory animals would actually occur in humans (Kimmel, C.A. and Kimmel, 1994). Thus, it cannot be assumed automatically that a growth deficit induced in a test animal will be expressed as such in a human. If sufficient information were available to make an informed judgment about similarity of effect between the test species and humans, growth or behavioral deficits might be judged to be reversible under some circumstances.

Table II-6. Major Component of the Reproductive Process Where Toxic Effects May Occur.

Developmental • Prenatal • Postnatal to Puberty	**Female Menstral or Estrous Cycle**
Gametogenesis • Oogenesis and Follicle Development • Ovulation • Spermatogenesis and Sperm Maturation	**Fertilization** **Gestation and Parturition Luctation** **Lactation** **Gametogenesis**

As indicated previously, all primordial follicles are formed in the ovaries of female mammals by or soon after birth, and the capability does not exist to replace them. In humans, premature depletion of the follicle population results in premature onset of menopause. Therefore, exposures that result in a reduced number of ovarian follicles initially or an accelerated rate of loss of follicles postpubertally are irreversible effects that can have severe consequences.

The reversibility of effects on spermatogenesis in males has been described previously in the section on Nature of Toxicity.

Female cycle normality

Estrous or menstrual cycle normality is dependent on normal ovarian function. Normal ovarian function is, in turn, dependent on the integrity of the hypothalamic and pituitary endocrine functions as well as ovarian paracrine control and a continuing supply of normal primordial and developing follicles. Thus, there are numerous sites and mechanisms through which cycle normality can be affected. They are too numerous and complex to consider here. Many of those effects can be reversible, but others, such as depletion of or unrepaired genetic damage to the follicle population would be irreversible. To judge the potential for recovery from disruption of the cycle, it would be necessary to know the underlying cause of the effect.

Fertility

On an even larger scale than with the female cycle, fertility of either a female or male requires the integration of multiple physiologic functions that are operating normally. These include adequate spermatogenesis and sexual behavior in males. In females, fertility requires ovulation and sexual receptivity as well as provision of an environment in the reproductive tract that is conducive to proper gamete transport, sperm capacitation and fertilization. Following fertilization, females must provide support for developing embryos and fetuses as well as normal parturition and adequate lactation.

Because fertility and pregnancy outcomes are apical end points, perturbation of any essential component can affect the final product adversely. Developmental effects have been considered previously. With respect to ability to achieve conception and deliver, infertility could be reversed in circumstances where the underlying cause is reversible. However, the cost of even temporary impaired fertility to human couples in terms of interference with family or career planning and psychological trauma should be considered, particularly if the source of the impairment and prognosis are unknown.

REVERSIBILITY AND RISK ASSESSMENT

Currently, potential to recover from an adverse effect is not considered in determining a reference dose (RfD) or reference concentration (RfC). It is important to decide whether potential for recovery should be cause for a lower level of concern in risk assessment or risk management. If the general concept of reduced concern for reversible effects is adopted, the mechanism for incorporation of that reduced concern into risk assessment and risk management must be determined.

The risk assessment paradigm for non-cancer health effects of environmental agents (National Research Council, 1983) provides two places to express reduced concern: in uncertainty factor

selection and in the risk characterization. Currently, the system used to assign uncertainty factors is designed primarily to account for missing information and is insensitive to fine adjustments. Information conveyed to risk managers in the risk characterization could influence the balance in decision-making. With consumer products or in occupational situations, reduced concern under some conditions of exposure might be conveyed in labels or warnings.

An issue that could be overriding is whether sufficient information will be available to determine that full recovery is likely in more than a small proportion of cases. Without such information, reversibility is a moot point.

REFERENCES

Chapin, R.E. and Ku, W.W. 1994. The reproductive toxicity of boric acid. *Envir. Health Perspect.* 102 (Suppl.7):211–214.

Chapin, R.E., Dutton, S.L., Ross, M.D., Swaisgood, R.R., and Lamb, J.C. 1985. The recovery of the testis over 8 weeks after short-term dosing with ethylene glycol monomethyl ether: Histology, cell-specific enzymes, and rete testis fluid protein. *Fund. Appl. Toxicol.* 5:515–525.

Cooke, P.S., Porcelli, J., and Hess, R.A. 1992. Induction of increased testis growth and sperm production in adult rats by neonatal administration of the goitrogen propylthiouracil (PTU): The critical period. *Biol. Reprod.* 46:146–154.

Foster, P.M.D., Creasy, D.M., Foster, J.R., Thomas, L.V., Cook, M.W., and Gangolli, S.D. 1983. Testicular toxicity of ethylene glycol monomethyl and monoethyl ethers in the rat. *Toxicol. Appl. Pharmacol.* 69:385–399.

Goldman, J.M., Cooper, R.L., Edwards, T.L., Rehnberg, G.L., McElroy, W.K., and Hein, J.F. 1991. Suppression of the luteinizing hormone surge by chlordimeform in ovariectomized, steroid-primed female rats. *Pharmacol. Toxicol.* 68:131–136.

Kimmel, C.A., and Kimmel, G.L. 1994. Risk assessment for developmental toxicity. In *Target Organ Toxicology Series: Developmental Toxicology*, ed. C. Kimmel and J. Buelke-Sam. New York: Raven Press, pp. 429–453.

Kimmel, G.L., Clegg, E.D., and Crisp, T.M. 1995. Reproductive toxicity testing: A risk assessment perspective. In *Reproductive Toxicology*, 2nd ed., ed. R.J. Witorsch. New York: Raven Press, pp. 75–98.

Manson, J.M. and Kang, Y.J. 1994. Test methods for assessing female reproductive and developmental toxicology. In *Principles and Methods of Toxicology*, 3rd ed., ed. A.W. Hayes. New York: Raven Press, pp. 989–1037.

National Research Council. 1983. *Risk Assessment in the Federal Government: Managing the Process*. Washington: National Academy Press.

Russell, L.D., Ettlin, R., Sinha Hikim, A.P., and Clegg, E.D. 1990. *Histological and Histopathological Evaluation of the Testis*. Clearwater, FL: Cache River Press.

Scialli, A.R. and Clegg, E.D. 1992. *Reversibility in Testicular Toxicity Assessment*. Boca Raton: CRC Press.

Sharpe, R.M. and Skakkebaek, N.E. 1993. Are estrogens involved in falling sperm counts and disorders of the male reproductive tract? *Lancet* 341:1392–1395.

Stoker, T.E., Goldman, J.M., Cooper, R.L., and McElroy, W.K. 1991. Influence of chlordimeform on alpha-adrenergic receptor-associated mechanisms of hormonal regulation in the rat: Pituitary and adrenocortical secretion. *Toxicol.* 69:257–268.

U.S. Environmental Protection Agency. 1991. Guidelines for Developmental Toxicity Risk Assessment. *Fed. Reg.* 56(234):63798–63826.

U.S. Environmental Protection Agency. 1994. Draft Guidelines for Reproductive Toxicity Risk Assessment. EPA/600/AP-94/001.

Zenick, H., Clegg, E.D., Perreault, S.D., Klinefelter, G.R., and Gray, L.E. 1994. Assessment of male reproductive toxicity: A risk assessment approach. In *Principles and Methods of Toxicology*, 3rd ed., ed. A.W. Hayes. New York: Raven Press, pp. 937–988.

IMMUNOTOXICITY AND RISK ASSESSMENT: EFFECT OF TEMPORAL FACTORS[2]

MaryJane K. Selgrade, Ph.D.
Healh Effects Research Laboratory
U.S. Environmental Protection Agency

SUMMARY

The immune system is composed of cells collectively best known as leukocytes. They are derived from bone marrow and are widely distributed throughout the body, which they defend against infections and certain tumors. Immune cells may also react to relatively innocuous foreign substances or against "self" components resulting in allergic or autoimmune disease, respectively. There is evidence that various chemicals can act as either immunosuppressants, allergens, or inducers of autoimmunity. A number of immune function tests have been developed to identify chemicals with immunosuppressive potential. However, immune suppression only poses a hazard in the presence of infectious agents or tumor cells. Temporal relationships between exposure to a compound and exposure to these agents have an impact on the risk associated with immune suppression. Three examples from animal research are presented to illustrate how temporal factors may affect the risk of infection: (1) Effects of immunosuppressants on natural killer cell activity and susceptibility to murine cytomegalovirus infection; (2) Effects of phosgene on alveolar macrophage phagocytic function and susceptibility to *Streptococcal zooepidemicus* infection in mice and rats; and (3) Effects of ultraviolet radiation on delayed-type hypersensitivity responses and susceptibility to *Leishmania major* and Mycobacterial infections.

INTRODUCTION

There are a number of resources that provide general background information on the physiology, biochemistry, and toxicology of the immune system (Roitt, 1988; Abbas et al., 1991; Dean et al., 1994; National Research Council, 1992; Dean and Murray, 1991; Descotes, 1986; U.S. Congress 1991). Briefly, the immune system defends the body against infectious agents (bacteria, viruses, fungi,

[2] Disclaimer: The research chapter has been reviewed by the Health Effects Research Laboratory, U.S. Environmental Protection Agency and approved for publication. Approval does not signify that the contents necessarily reflects the views and policies of the agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

parasites), which are ubiquitous in our environment, and also appears to play a role in defending the body against certain tumor cells, which may arise spontaneously or as the result of environmental insults (viral, radiation, and possibly chemical). As such, a properly functioning immune system is essential to good health. Individuals with compromised immune systems as a result of congenital defects, disease (AIDS, leukemia), or drug therapies are more susceptible to infectious diseases and certain types of cancer, the consequences of which can be life-threatening (Abbas et al., 1991; Penn, 1985, 1993; Pinching, 1984; Ammann, 1984). On the other hand, the immune system may react to foreign substances that would otherwise be relatively innocuous, such as pollen or house dust, resulting in allergic reactions or with "self" components resulting in autoimmune disease. There is evidence that assorted chemical exposures can alter the immune system either as immunosuppressants, allergens, or inducers of autoimmunity and as a result produce adverse health effects. This paper will be limited to risk assessment issues associated with immune suppression, specifically the impact that the temporal relationship between exposure to an immunotoxicant and exposure to an infectious agent may have on health effects.

General Background on the Immune System and Immunotoxicity Testing

The immune system is composed of several tissues and cell types (Figure II-7). The cells are collectively best known as white blood cells or leukocytes and include macrophages and monocytes, neutrophils, basophils, mast cells, eosinophils, natural killer (NK) cells, T and B lymphocytes, and plasma cells (B cells which produce antibodies). They are derived from stem cells in bone marrow (or fetal liver), circulate in blood and lymph and are widely distributed in lymphoid and other tissues throughout the body. The maturation and differentiation of these cells are controlled by soluble factors, such as cytokines, some of which are produced by cells outside the immune system (e.g., epithelial cells).

Exposure to a foreign antigen triggers a series of complex, but highly integrated responses. One of the challenges in devising test guidelines for immunotoxicity has been to determine the most appropriate tests for adequately assessing the integrity of the immune system. Because of the complexity of the immune system, tiered testing approaches have frequently been employed using either mice or rats (Dean and Vos, 1986; Vos and van Loveren, 1987; Exon et al., 1990; Luster et al., 1988; Sjoblad, 1988). In some cases the first level of the tier relies solely on non-functional endpoints including changes in the weight of thymus and other lymphoid organs, histopathology of these organs, and/or differential blood cell counts (Vos and van Loveren, 1987; Vos et al., 1994). However, an

analysis of a large chemical database in mice indicated that these endpoints by themselves may not be very accurate in predicting changes in immune function or alterations in susceptibility to challenge with infectious agents or tumor cells, particularly following low dose exposures (Luster et al., 1992, 1993). Hence, the first tier of tests often includes a limited number of functional assays designed to assess as efficiently as possible the integrity of the major components of the immune system.

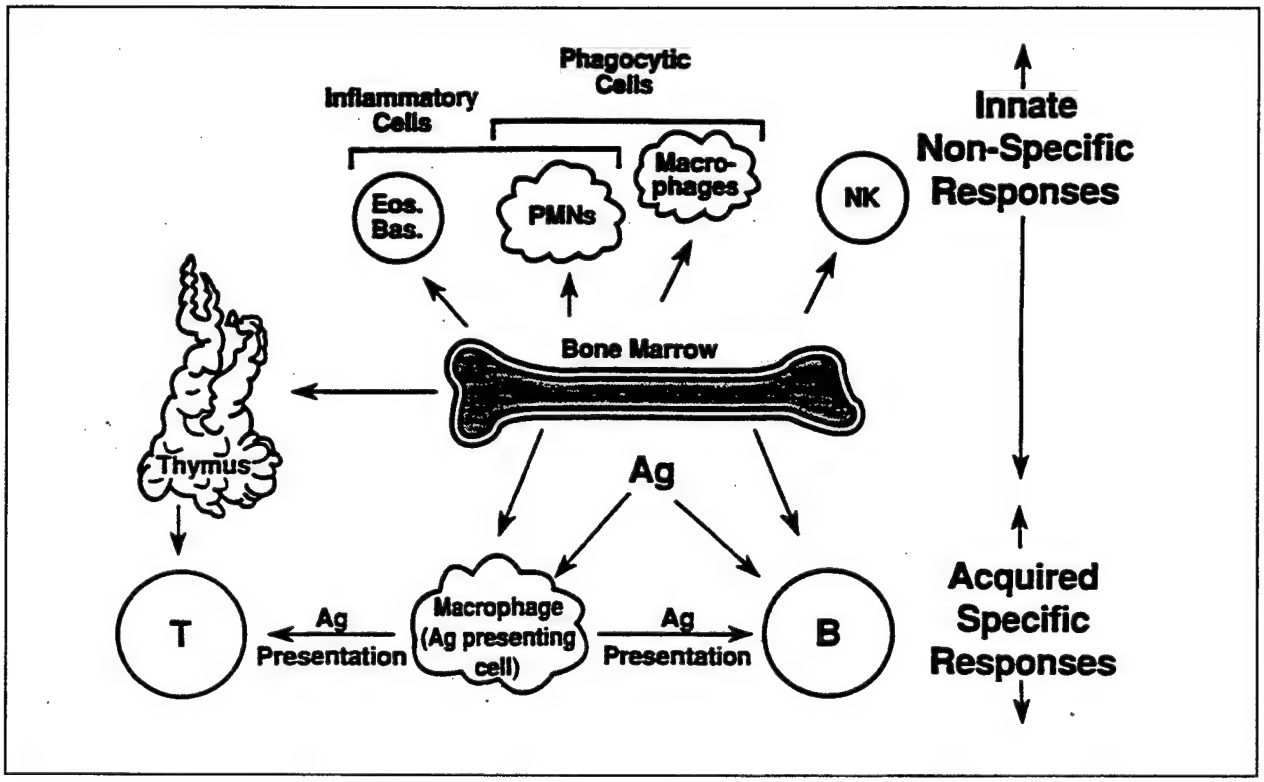


Figure II-7. The Immune System. The immune system is composed of primary lymphoid organs (bone marrow and thymus), secondary lymphoid organs (not shown, but including spleen, lymph nodes, etc.), and several cell types including macrophages, natural killer cells (NK), polymorphonuclear leukocytes (PMNs), eosinophils (Eos), basophils (Bas), T and B lymphocytes.

The National Toxicology Program (NTP) testing battery (Luster et al., 1988) and the Environmental Protection Agency's (EPA) immunotoxicity testing guidelines for biochemical pest control agents (Subdivision M) (Sjoblad, 1988) are similar and representative of this approach (Table II-7). Tier 1 includes functional assays indicative of (1) antibody-mediated responses, (2) T-cell-mediated responses, and (3) NK cell activity. The EPA also includes a test for macrophage function. Tier II represents a more in-depth evaluation of these responses and also includes assays that evaluate

the ability of animals to actually resist challenge with infectious agents or transplantable tumor cells. Selection of appropriate tier II tests to be performed are usually based, at least in part, on results of tier I. For many chemicals only tier I data are available. Recent data from the NTP indicate that, taken together, tier I tests are predictive for both immunosuppressive effects (Luster et al., 1992) and effects on resistance to challenge with an infectious agent or tumor cells (Luster et al., 1993). Several tests, including the antibody-forming cell assay, lymphocyte cell surface marker analysis, and NK cell activity, alone or in combination with one other test, were very predictive of immunotoxicity, suggesting that a good indication of immunotoxicity and susceptibility to infection and tumors in animals may be obtained using a minimal number of selected tests Luster et al., 1992, 1993).

Table II-7. Parameters For Immunotoxicity Testing in Rodents*.

Parameter	Procedure/Immunoassay
Tier I (Screen)	
Immunopathology	Hematology; Weights: body, spleen, thymus
	Cellularity: spleen
	Histology: Spleen, thymus, lymph node
Humoral immunity	IgM plaque-forming cell (PFC) response (to SRBC)
	Lymphoproliferative response: B cell mitogens
Cell-mediated immunity	Lymphoproliferative response: T cell mitogens
	Allogeneic mixed leukocyte response (MLR)
Nonspecific immunity	Natural killer (NK) cell activity
Tier II (Comprehensive)	
Immunopathology	Bone marrow cellularity, histology, and stem cell activity (e.g., CFC assays)
	Quantitation of B and T cell numbers
Humoral immunity	IgG PFC to SRBC
	IgM PFC to T cell-independent antigen (e.g., TNP-LPS)
Cell-mediated immunity	Cytotoxic T lymphocyte (CTL) cytolysis
	Delayed hypersensitivity response (DHR)
Nonspecific immunity	Macrophage: quantitation (resident peritoneal cells)
	Macrophage: function (phagocytosis, bactericidal and tumoricidal activity)
	Neutrophil: function (phagocytosis and bactericidal activity)
Host resistance models	Response to challenge with infectious agent or tumor cells

*Adapted from Luster et al., 1988.

Risk Assessment and Temporal Factors

Two of the most problematic uncertainties which affect the interpretation of results from immunotoxicity tests in terms of potential risk to human health are the need to extrapolate effects across levels of biologic organization and the need to extrapolate from animal data to human health effects. The first issue involves uncertainties associated with establishing a quantitative relationship between changes in individual immune function tests (at the cellular level) and altered resistance to infectious or neoplastic disease (at the population level). These issues are not unique to immunotoxicity, but apply generally to all noncancer endpoints. Possible methods for handling these uncertainties have been reviewed elsewhere (Selgrade et al., 1994).

The role that temporal factors play in the outcome of immunotoxic events represents another issue that must be considered in the interpretation of such data. With respect to immunosuppression, two components are required in order to have a health hazard: (1) exposure to an immunosuppressive compound, and (2) exposure to an infectious agent or neoplastic cell. Obviously the longer an immunosuppressive state is maintained, the greater the odds are that an infectious agent will be encountered and an adverse effect will ensue. However, a single ill-timed exposure may have serious long-term consequences if it happens to coincide with exposure to an infectious agent that may subsequently damage target organs more severely than in an immunocompetent host. Since there is fairly rapid turnover of immune cells, chronic exposure is usually required to maintain immune suppression over long periods of time (thus increasing the odds that an adverse effect will occur as a result of challenge with an infectious agent). However, long-term immune suppression could result from an acute exposure, if a chemical were to target stem cells, if exposure occurred during periods of critical development (e.g. thymic selection in perinatal development), or if exposure resulted in immunologic tolerance towards particular antigens (as described below for ultraviolet radiation [UVR]).

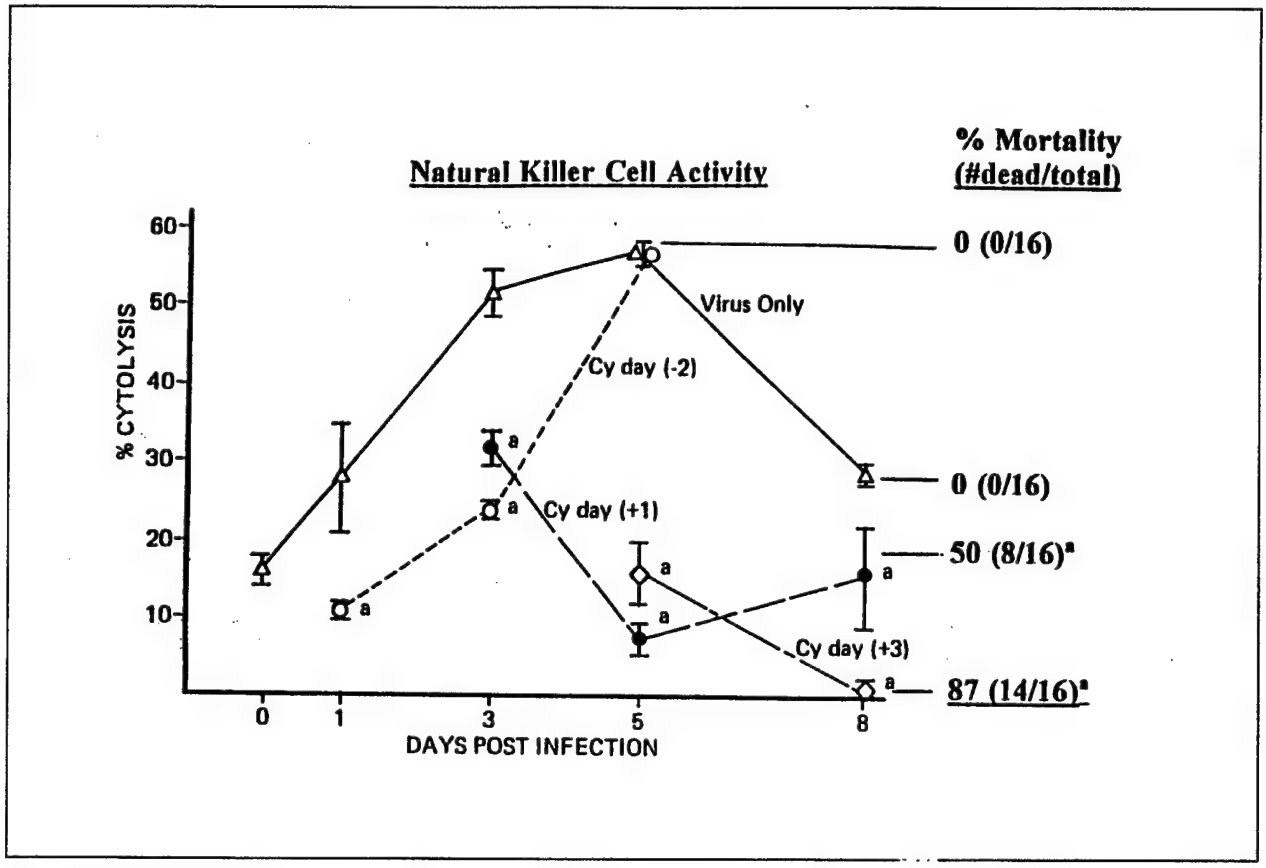
In order to include a consideration of temporal factors in a risk assessment, it is necessary to have an understanding of the mechanisms associated with immunosuppression, the types of host responses affected by a particular chemical, and the role that these responses play in defending the host during the various stages of infection. Animal data from three types of experiments including the effects of immunosuppressants on NK cell activity and susceptibility to murine cytomegalovirus (MCMV), the effects of acute and chronic phosgene exposure on alveolar macrophage function and susceptibility to streptococcal infection in mice and rats, and the effects of ultraviolet radiation on delayed-type

hypersensitivity and susceptibility to several infections, are used to illustrate various ways in which temporal factors may affect the outcome of exposure to immunosuppressive chemicals and infectious agents.

Effects of Immunosuppressants on Natural Killer Cell Activity and MCMV Infection

Murine cytomegalovirus (MCMV) is a well-established animal model for human cytomegalovirus infection (Osborn, 1982), a herpesvirus and opportunistic infection that causes life-threatening problems in transplant patients and individuals suffering from AIDS or other immunosuppressive disorders (Betts and Hanshaw, 1977; Jacobson and Mills, 1988; Preiksaitis, 1989). The earliest host response to MCMV infection, interferon production, is detectable in serum 6-h post infection (p.i.), disappears by 6 days p.i. (Catignani et al., 1989) and is thought to cause augmentation of NK cell activity, which is apparent 1 day p.i. and wanes approximately 8 days p.i. (Selgrade et al., 1982). These two responses appear to play crucial roles in recovery from acute MCMV infection (Bancroft et al., 1981; Shellam et al., 1982) of spleen, liver, and lung, where virus titers peak during the first week of infection and wane during the second week of infection (Selgrade et al., 1984). Several xenobiotic compounds given by a variety of routes have been shown to suppress virus-augmented NK cell activity and enhance susceptibility to MCMV (reviewed by Selgrade et al., 1992). Suppression of NK activity must occur during the first week of infection. Figure II-8 illustrates how the temporal relationship between administration of the immunosuppressant cyclophosphamide (150 mg/kg), and infection can affect the disease outcome. In this study cyclophosphamide suppressed virus-augmented NK cell activity; however, recovery was fairly rapid. Hence, when a single dose of cyclophosphamide was given 2 days prior to infection there was no effect on mortality due to infection because, while NK activity was initially suppressed, by 5 days p.i. NK activity levels were the same as in untreated mice. Natural killer activity in these mice apparently recovered in time to prevent mortality. However, if cyclophosphamide was given 1 or 3 days postinfection, NK cell activity was suppressed during a critical time in infection, and mortality due to infection was enhanced. It should be noted that for another immunosuppressive compound, 7,12-dimethylbenz[*a*]anthracene (DMBA), suppression of virus-augmented NK cell activity lasted for a much longer period of time. For example, in mice treated 2 weeks prior to infection with DMBA suppression of virus-augmented NK cell activity was observed and enhanced mortality due to infection occurred (Selgrade et al., 1988). In order to demonstrate enhanced infection in this infectivity model, the temporal relationship between exposure to the immunotoxicant and infection must be such that virus-augmented NK cell activity is suppressed during

the first week of infection. The timing necessary to achieve this effect varies depending on the immunotoxicant.



[a] = significantly different from virus only ($p < 0.05$).

Figure II-8. Effect of the temporal relationship between administration of cyclophosphamide and MCMV infection on suppression of natural killer cell activity and mortality due to infection. Mice were treated with 150 mg/kg 2 days prior to infection (O), or 1 day (⊕), or 3 days (◊) postinfection and compared to vehicle treated controls (_). Natural killer cell activity was assessed as % cytotoxicity on various days post- infection, or mice were observed for mortality (adapted from Selgrade et al., 1982).

Effect of Phosgene on Alveolar Macrophage Function and Susceptibility to Streptococcal Infection

Phosgene is a toxic gas widely used in industrial processes. A sensitive endpoint for acute phosgene toxicity in both mice and rats is decreased resistance to challenge with the bacteria *Streptococcus zooepidemicus* (Selgrade et al., 1989; Yang et al., 1994). This effect is attributed to impaired alveolar macrophage (AM) phagocytosis of bacteria (reviewed by Selgrade and Gilmour, 1994). Enhanced susceptibility to infection in mice is measured by enhanced mortality following

infection. In rats enhanced susceptibility is measured by delayed clearance of bacteria from the lung and by increased numbers of polymorphonuclear leukocytes in lung lavage fluid (Gilmour and Selgrade, 1993). In both cases enhanced susceptibility was observed when infection occurred shortly after a single phosgene exposure (3 h to 0.025 ppm or 8 h to 0.01 ppm for mice; 6 h to 0.1 ppm for rats) (Selgrade et al., 1989; Yang et al., 1994). If infection was delayed until 24 h after exposure, susceptibility to infection was not enhanced, suggesting that, as with NK cells in the previous example, AM function recovers fairly rapidly following phosgene exposure. It should be noted that following inhalation exposure to $NiCl_2$ and $NiSO_4$, enhanced mortality in mice was actually greater when infection was delayed until 24 h after exposure (Adkins et al., 1979), again demonstrating that temporal relationships critical to enhanced infection may be different for different immunotoxicants.

Because recovery of AM function following a single phosgene exposure is rapid, the opportunity for enhanced infection is short-lived. However, recent studies demonstrated enhanced susceptibility to infection in rats exposed to 0.1 or 0.2 ppm phosgene, for 6 h/day, 5 days per week, for 4 or 12 weeks, and infected shortly after the last exposure (Selgrade et al., manuscript in preparation). The enhanced susceptibility was equal to that observed after a single acute exposure suggesting that adaptation did not occur. Presumably increased risk of infection existed for a short period of time (4–6 h) following each phosgene exposure in this chronic study, greatly expanding the opportunity for enhanced infection as compared to a single exposure. From this example it is clear that recovery time, exposure regimen, and potential adaptation are all factors which can influence the outcome of infection.

Effects of Ultraviolet Radiation on Delayed-type Hypersensitivity and Susceptibility to Infection

Mice exposed to minimal erythemal doses of ultraviolet radiation and subsequently challenged with a contact sensitizer either at the site of irradiation or at a distant site, failed to develop a delayed-type hypersensitivity (DTH) response as measured by the mouse ear swelling test (Toews et al., 1980; Noonan et al., 1981). In similar studies irradiated mice also failed to develop DTH responses to protein antigens (Ullrich et al, 1986; Jeevan & Kripke, 1990). The failure to respond appeared to be due to the development of antigen specific suppressor T cells, which rendered the mouse tolerant to that particular antigen. This concept was supported by the demonstration that mice irradiated and sensitized at one time point could not be sensitized with the same chemical through unirradiated skin 14 days later and that antigen-specific unresponsiveness could be adoptively transferred to naive mice by spleen and lymph node cells obtained from mice skin-painted through irradiated skin (Elmets et al., 1983). Hence a single exposure to UVR produced long-term suppression to a given antigen. Current thinking

(reviewed by UNEP/WHO/ICNIRP, 1994) is that UVR causes the release of mediators (cytokines) from the skin, which alter the antigen presenting capability of Langerhans cells as well as antigen presenting cells at other sites. It is conceivable that other inflammatory agents may trigger similar cytokine responses (Araneo et al., 1989). The alteration in antigen presentation results in the development of cells with suppressor activity. It may be that these "suppressor" cells are T helper (TH)2 cells. The net effect is failure to activate TH1 cells and suppression of DTH responses thought to play an important role in host defenses against certain types of tumors and microbial infections. The immune suppression is antigen specific, (i.e., only responses to antigens administered within 7 days after irradiation are affected) and long lasting (at least 3 months).

The consequences of UVR-induced immune suppression for several types of infectious disease have been explored using murine models. The tails of mice were irradiated with suberythemal doses (0.06–6 kJ/m^2) of UVR, 3 times per week, for 1 month and infected with *Leishmania major* intradermally through the irradiated surface of the tail 24 h after the first UVR exposure. In UVR-treated mice, the DTH response to *L. major* antigens 2- and 6-weeks postinfection was suppressed. The lesion due to infection at the site of injection was greatly reduced in mice exposed to UVR, but the number of organisms recovered from skin at the injection site was actually greater in UVR-exposed mice, suggesting that the immune response contributed to the lesion (Giannini, 1986a). In addition, a larger number of parasites was observed in the draining lymph node in UVR-infected mice (Giannini, 1987, 1992). Finally, mice infected through irradiated skin failed to develop protective immunity such that lesions following reinfection at an unirradiated site were significantly larger when compared to lesions of previously infected but unirradiated mice (Giannini, 1986b).

In another infectious disease model, mice were irradiated on the shaved back and injected with BCG (the vaccine strain for tuberculosis) subcutaneously in the footpad. Mice exposed from 1 to 15 times (3 times per week for up to 5 weeks) to one minimal erythemal dose (2.25 kJ/m^2) showed significant suppression in their DTH response to tuberculin and increased numbers of live bacteria in the spleen and lymph node compared to unirradiated controls (Jeevan & Kripke, 1990). However, when exposures were continued beyond five weeks, the DTH response recovered and mice challenged with bacteria at that point did not exhibit increased numbers of organisms in the spleen and lymph node, suggesting that adaptation to UVR may eventually occur.

In a similar model, mice treated with a single high UVR dose (45 kJ/m^2) 3 days before infection with *Mycobacterium lepraemurium* exhibited significant suppression of DTH responses to

mycobacterial antigen 3 and 6 months after infection and had significantly more bacteria in the infected footpad, lymph node, and spleen 3–6 months postinfection (Jeevan et al., 1992). This high dose also reduced the median survival time of mice infected intravenously. With a lower exposure dose (2.3 kJ/m^2), 50% suppression of the DTH response to mycobacterial antigen was observed three months postinfection, and increased numbers of bacteria were observed in the footpad, spleen, and lymph node of mice exposed to UVR doses greater than or equal to 5.6 kJ/m^2.

It is clear from the foregoing example that UV-induced immune suppression can have consequences that are manifested weeks and/or months after a single or repeated exposure. It is also evident that after a number of exposures adaptation may occur. These results indicate that several temporal factors may influence the consequences of UVR-induced immune suppression.

SUMMARY

In summary, immune suppression is only a hazard when it is accompanied by exposure to infectious agents or tumor cells. Since infectious agents are ubiquitous in our environment, the likelihood that an immunosuppressed individual will encounter these agents is great, and the more prolonged the immune suppression the greater the risk becomes. The effect that the temporal relationship between exposure to an immunotoxic and exposure to an infectious agent may have on the risk of disease will differ with different toxicants and infectious agents. As the examples above illustrate, the types of immune responses affected by a chemical and their importance to defense against a particular infectious agent, the recovery time of the immune response following chemical exposure, the length of exposure and the potential for adaptation of the immune response to the toxicant, or development of immunologic tolerance to a particular agent may all influence the risk of exposure. Obviously, the best estimates of risk can be made when the mechanisms associated with immunosuppression are known and the above mentioned variables have been defined.

REFERENCES

Abbas, A.K., Lichtman, A.H., and Pober, J.S. 1991. *Cellular and Molecular Immunology*. Philadelphia: W.B. Saunders.

Adkins, B., Richards, J.H., and Gardner, D.E. 1979. Enhancement of experimental respiratory infection following nickel inhalation. *Environ. Res.* 20:33–42.

Ammann, A.J. 1984. Immunodeficiency Disease. In: *Basic and Clinical Immunology*, 5th Edition, eds. D.P. Stites, J.D. Stobo, H.H. Fundenburg, and J.V. Wells, pp. 384–422. Los Altos, CA: Lange Medical Publishers.

Araneo, B.A., Dowell, T., Moon, H.B., and Daynes, R.A. 1989. Regulation of murine lymphokine production *in vivo*. Ultraviolet radiation exposure depresses Il-2 and enhances IL-4 production by T cells through an Il-1 dependent mechanism. *J. Immunol.* 143: 1737–1744.

Bancroft, G.J., Shellam, G.R., and Chalmer, J. 1981. Genetic influences on the augmentation of natural killer (NK) cells during murine cytomegalovirus infection. Correlation with patterns of resistance. *J.Immun.* 126:988–994.

Betts, R.F. and Hanshaw, J.B. 1977. Cytomegalovirus (CMV) in the compromised host(s). *Ann. Rev. Med.* 28:103–110.

Catignani, J.C., Menache, M.G., and Selgrade, M.J.K. 1989. Increased susceptibility to pentobarbital following mouse cytomegalovirus infection: Relative roles of viral-induced interferon and viral infection of the liver. *J. Biochem. Toxicol.* 4:1–9.

Dean, J.H. and Murray, M.J. 1991. Toxic Responses of the Immune System, in: *Casarett and Doull's Toxicology: The Basic Science of Poisons*, M.O. Amdur, J. Doull, and C.D. Klaassen, eds., pp. 282–333, New York: McGraw-Hill.

Dean, J.H. and Vos, J.G. 1986. An Introduction to Immunotoxicology Assessment. In: *Immunotoxicology of Drugs and Chemicals*, J. Descotes, ed., pp. 3–18. New York: Elsevier.

Dean, J.H., Luster, M.I., Munson, A.E., and Kimber, I. 1994. *Immunotoxicology and Immunopharmacology*. New York: Raven Press.

Dean, J.H. and Murray, M.J. 1991. Toxic Responses of the Immune System. *In Casarett and Doull's Toxicology: The Basic Science of Poisons*, M.O. Amdur, J Doull, and C.D. Klaassen, eds., pp. 282–333. New York: McGraw-Hill.

Descotes, J. 1986. *Immunotoxicology of Drugs and Chemicals*. New York: Elsevier.

Elmets, C.A., Bergstresser, P.R., Tigelaar, R.E., Wood, P.J., and Streilein, J.W. 1983. Analysis of the mechanism of unresponsiveness produced by haptens painted on skin exposed to low dose ultraviolet radiation. *J. Exp. Med.* 158:781–794.

Exon, J.H., Bussiere, J.L., and Mather, G.G. 1990. Immunotoxicity Testing in the Rat: An Improved Multiple Assay Model. *Int. J. Immunopharmac.* 12:699–701.

Giannini, S.H. 1986a. Suppression of pathogenesis in cutaneous leishmaniasis by UV irradiation. *Infect. Immun.* 51:838–843.

Giannini, S.H. 1986b. Effects of UV-B on infectious disease. In: *Effects of Changes in Stratospheric Ozone and Global Climate*, J.G. Titus, eds., pp. 101–112. Washington, DC: U.S. Environmental Protection Agency.

Giannini, S.H. 1987. Abrogation of skin lesions in cutaneous leishmaniasis by ultraviolet irradiation. In: *Leishmaniasis: The First Centenary (1885–1985) New Strategies for Control.* NATO ASI series A: Life Sciences, D.T. Hart, ed., 677–684. London: Plenum.

Giannini, S.H. 1992. Effects of ultraviolet B irradiation on cutaneous leishmaniasis. *Parasit. Today* 8:44–48.

Gilmour, M.I. and Selgrade, M.J.K. 1993. A comparison of the pulmonary defenses against Streptococcal infection in rats and mice following O_3 exposure: A possible of disease resistance in rats. *Toxicol. Appl. Pharmacol.* 123:211–218.

Jacobson, M.A. and Mills, J. 1988. Serious cytomegalovirus disease in the acquired immunodeficiency syndrome (AIDS. CLinical findings, diagnosis, and treatment). *Ann. Intern. Med.* 108:585–594.

Jeevan, A. and Kripke, M.L. 1990. Alteration of the immune response to *Mycobacterium bovis* BCG in mice exposed chonically to low doses of UV radiation. *Cell. Immunol.* 130:32–41.

Jeevan, A., Gilliam, K., Heard, H., and Kripke, M.L. 1992. Effects of ultraviolet radiation on the pathogenesis of *Mycobacterium lepraemurium* infection in mice. *Exp. Dermatol.* 1:152–160.

Luster, M.I., Munson, A.E., Thomas, P.T., Holsapple, M.P., Fenters, J.D., White, K.L., Jr., Lauer, L.D., Germolec, D.R., Rosenthal, G.J., and Dean, J.H. 1988. Methods Evaluation: Development of a Testing Battery to Assess Chemical-Induced Immunotoxicity: National Toxicology Program's Guideline for Immunotoxicity Evaluation in Mice. *Fundam. Appl. Toxicol.* 10:2–19.

Luster, M.I., Portier, C., Pait, D.G., White, K.L., Jr., Gennings, C., Munson, A.E., and Rosenthal, G.J. 1992. Risk assessment in immunotoxicology I. Sensitivity and predictability of immune tests. *Fundam. Appl. Toxicol.* 18:200–210.

Luster, M.I., Portier, C., Pait, D.G., Rosenthal, G.J. Germolec. D.R., Corsini, E., Blaylock, B.L., Pollock, P., Kouchi, Y., Craig, W., White, D.L., Munson, A.E., and Comment, C.E. 1993. Risk Assessment in Immunotoxicology II. Relationships Between Immune and Host Resistance Tests. *Fundam. Appl. Toxicol.* 21:71–82.

National Research Council. 1992. *Biologic Markers in Immunotoxicology.* Washington, DC: National Academy Press.

Noonan, F.P., Kripke, M.L., Pedersen, G.M., and Green, M.I. 1981. Suppression of contact hypersensitivity in mice by ultraviolet irradiation is associated with defective antigen presentation. *Immunology* 43:527–533.

Osborn, J.E. 1982. CMV-herpesviruses of mice. In *The Mouse in Biomedical Research*, H.L. Foster, J.G. Fox, and J.D. Small, eds, 2:267–292. New York: Academic Press.

Penn, I. 1985. Neoplastic Consequences of Immunosuppression. In *Immunotoxicology and Immunopharmacology*, J.J. Dean, M.I. Luster, A.E. Munson, and H. Amos, eds., pp. 367–380. New York: Raven Press.

Penn, I. 1993. Tumors After Renal and Cardiac Transplantation. *Hematol. Oncol. Clin. North Am.* 7:431–435.

Pinching, A.J. 1984. The Acquired Immune Deficiency Syndrome. *Clin. Exp. Immunol.* 56:1–13.

Preiksaitis, J.K. 1989. Cytomegalovirus infection in transplant recipients. *Immunol. Allergy Clinics of N. Amer.* 9:137–151.

Roitt, I.M. 1988. *Essential Immunology.* Cambridge, MA: Blackwell Scientific Publications.

Selgrade, M.J.K., Daniels, M.J., Hu, P.C., Miller, F.J., and Graham, J.A. 1982. Effects of immunosuppression with cyclophosphamide on acute murine cytomegalovirus infection and virus-augmented natural killer cell activity. *Infect. Immun.* 38:1046–1055.

Selgrade, M.J.K., Collier, A.M., Saxton, L., Daniels, M.J., and Graham, J.A. 1984. Comparison of the pathogenesis of murine cytomegalovirus in lung and liver following intraperitoneal or intratracheal infection. *J. Gen. Virol.* 65:515–523.

Selgrade, M.J.K., Daniels, M.J., Burleson, G.R., Lauer, L.D., and Dean, J H. 1988. Effects of 7,12-dimethylbenz[*a*]anthracene, benzo[*a*] pyrene, and cyclosporin A on murinbe cytomegalovirus infection: Studies of resistance mechanisms. *Int. J. Immunopharmac.* 10:811–818.

Selgrade, M.J.K., Starnes, D.M., Illing, J.W., Daniels, M.J., and Graham, J.A. 1989. Effects of phosgene exposure on bacterial viral and neoplastic lung disease susceptibility in mice. *Inhalation Toxicol.* 1:243–259.

Selgrade, M.J.K., Daniels, M.J., and Dean, J.H. 1992. Correlation between chemical suppression of natural killer cell activity in mice and susceptibility to cytomegalovirus: Rationale for applying murine cytomegalovirus as a host resistance model and for interpreting immunotoxicity testing in terms of risk of disease. *J. Toxicol. Environ. Health* 37:123–137.

Selgrade, M.J.K. and Gilmour, M.I. 1994. Effects of gaseous air pollutants on immune responses and susceptibility to infectious and allergic disease. In: *Immunotoxicology and Immunopharmacology,* 2nd edition, J.H. Dean, M.I. Luster, A.E. Munson, and I. Kimber, 2, pp. 395–411. New York: Raven Press.

Selgrade, M.J.K., Cooper, K.D., Devlin, R.B., van Loveren, H., Biagini, R.E., and Luster, M.I. 1994. Immunotoxicity-bridging the gap between animal research and human health effects. *Fund. Appl. Toxicol.* submitted.

Shellam, G.R., Grundy, J.E., and Allan, J.E. 1982. The role of natural killer cells and interferon in resistance to murine cytomegalovirus. In: *NK Cells and Other Natural Effector Cells,* R.B. Herbermann, ed., pp. 1451–1457. New York: Academic Press.

Sjoblad, R.D., 1988. Potential Future Requirements for Immunotoxicology Testing of Pesticides. *Toxicol. Indust. Health* 4:391–394.

Toews, G.B., Bergstresser, P.R., and Streilein, J.W. 1980. Epidermal langerhans cell density determines whether contact hypersensitivity or unresponsiveness follows skin painting with DNFB. *J. Immunol.* 124:445–453.

Ullrich, S.E., Azizi, E., and Kripke, M.L. 1986. Suppression of the induction of delayed-type hypersensitivity reactions in mice by a single exposure to ultraviolet radiation. *Photochem. Photobiol.* 43: 633–638.

UNEP/WHO/ICNIRP. 1994. *Environmental Health Criteria Ultraviolet Radiation.* In press, Geneva: World Health Organization.

U.S. Congress, Office of Technology Assessment. 1991.

Identifying & Controlling Immunotoxic. Washington, DC: U.S. Government Printing Office.

Vos, J.G. and van Loveren, H. 1987. Immunotoxicity Testing in the Rat. In *Environmental Chemical Exposures and Immune System Integrity, Vol XIII, Advances in Modern Environmental Toxicology,* E.J. Burger, R.G. Tardiff and J.A. Ballanti, eds., pp. 167–180. Princeton, NJ: Princeton Scientific.

Vos, J.G., Smialowicz, R.J., and Van Loveren, H. 1994. Animal Models for Assessment. In: *Immunotoxicology and Immunopharmacology*, J.H. Dean, M.I. Luster, A.E. Munson, and I. Kimber, eds., New York: Raven Press.

Yang, Y.G., Gilmour, M.I., Lange, R., Burleson, G.R., and Selgrade, M.J.K. 1994. Effects of acute exposure to phosgene on pulmonary host defenses and resistance to infection. *Inhalation Toxicol.* in Press.

TEMPORAL FACTORS OF EXPOSURE IN IDENTIFYING HAZARDS: NEUROTOXICITY

Suzanne B. McMaster
Health Effects Research Laboratory
U.S. Environmental Protection Agency

ABSTRACT

Neurotoxicity is commonly viewed as one of the most complex endpoints of toxicity. Effects of toxic agents on the nervous system can be manifest in many ways. These include biochemical, structural and functional changes. Many different measures of effect can be obtained within each of these categories.

An understanding of the temporal relationship between measurable signs of neurotoxicity and exposure to a chemical agent is critical to the accurate assessment of hazard and risk. Effects may be seen immediately or may require a period of elapsed time or some biological trigger to develop. They may be permanent or may appear to be fleeting, intermittent or reversible. Delayed effects are often difficult to associate with a specific chemical exposure. Although some measures of neurotoxicity are irreversible, others are rapidly reversible and therefore more difficult to detect.

In agreement with the basic concepts of toxicology, irreversible effects on measures of neurotoxicity are generally considered adverse. However, when dealing with apparently reversible effects, the neurotoxicologist often departs from the basic principles of general toxicology. Recovery of function following exposure to a neurotoxic chemical and a period of evident toxicity does not always indicate that the neurotoxicity is reversible. Rather, recovery of function may reflect the inherent redundancy of the nervous system and result from the activation of a reserve capacity.

Activation of reserve capacity provides the ultimate example of delayed neurotoxicity and underscores the importance of temporal relationships between exposure and neurotoxicity. An individual who develops some form of neurotoxicity as a result of exposure early in life may appear to recover fully and function normally for many years. Late in life, however, when the nervous system would normally call upon its reserve capacity to compensate for the natural effects of aging, the

individual exposed earlier may find the reserve capacity to be diminished severely. Premature aging or rapid deterioration of function may result.

Neurotoxicity is commonly viewed as one of the most complex endpoints of toxicity. This complexity impacts on all aspects of neurotoxicity risk assessment, beginning with the hazard identification process. In order to begin to identify hazard for any endpoint of toxicity, it is necessary to consider three seemingly simple factors. The first is how the endpoint will manifest itself; the second, how it can be measured; and the third, the timepoint at which it should be evaluated. The latter of these factors, the temporal relationship between exposure to a potentially neurotoxic substance and identification of any resulting effect is the main topic of this paper. The field of neurotoxicity has made great progress toward understanding the workings of the nervous system and has assembled a comprehensive set of tools with which to examine this system. Perhaps the next step is the general recognition of the temporal aspect of neurotoxicity and consideration of that factor into risk assessment approaches.

NEUROTOXICITY

Neurotoxicity is defined as any adverse change in the structure, chemistry or function of the central and/or peripheral nervous system following exposure to a chemical, physical or biological agent (OTA, 1990). Adverse effects include unwanted effects and any alteration from baseline that diminishes an organism's ability to survive, reproduce or adapt to the environment. Functional or structural changes in the nervous system that (1) prevent the nervous system from functioning fully or (2) require compensatory changes to maintain full function are also included in the definition of adverse neurotoxicological effects.

GENERAL PRINCIPLES

General principles to guide the evaluation of neurotoxicity have been developed (FCCSET, 1993). They include:

- While some sub-populations may be more susceptible to certain neurotoxicants, neurotoxicity can occur at any time in the life span of an organism.
- Neurotoxicity can result from either direct or indirect action. Many neurotoxicants act directly on the nervous system. Other neurotoxicants produce states such as hypoxia which result in nervous system damage.
- Neurotoxicity can result from a single exposure. The effects may be apparent immediately or they may develop over time.

- Repeated exposure to a neurotoxicant may result in progressive alteration in the nervous system.

- Effects may be irreversible or reversible. Consideration should be given to possible underlying compensatory mechanisms when effects appear to be reversible.

- Latent or residual effects can sometimes be demonstrated following appropriate environmental or pharmacological challenge.

KEY ISSUES

Three key issues contribute to the difficulty of neurotoxicity hazard identification: (1) the complexity of the nervous system; (2) the identification of reversibility and compensation; and (3) the potential delay between exposure and effect. A fourth issue, the temporal relationship between exposure and effect is perhaps recognized the least. It is, however, a critically important issue. If overlooked, this relationship can result in a failure to identify neurotoxicity in otherwise well conducted hazard assessments.

Complexity

The nervous system is comprised of the brain, spinal cord and a complex network of nerve processes, transmitters, hormones, receptors and channels. Manifestation of neurotoxicity can be observed at multiple levels within an organism as described below. The complexity of the nervous system gives rise to an equally complex array of mechanisms of toxicity and potential effects. For accurate hazard identification, it is necessary to evaluate, directly or indirectly, the effect of a potentially neurotoxic agent at all levels of nervous system organization. Observation of an effect that is: (1) within the scope of the definitions of neurotoxicity; (2) adverse as described above; and (3) occurring at any level of nervous system organization is sufficient evidence of neurotoxicity for purposes of hazard identification.

Neurotoxic effects can be described at multiple organizational levels. The most commonly used descriptive levels are neurochemical, structural, neurophysiological and behavioral. At the neurochemical level, exposure to a neurotoxicant can inhibit protein or neurotransmitter synthesis, alter flow of ions across cellular membranes or prevent release of a neurotransmitter from nerve terminals. Structural effects of a neurotoxicant can include damage to a neuron, its axon or the myelin sheath. In addition, gross morphological changes such as increases or decreases in brain weight (absolute or relative to body weight), discoloration, hemorrhage or obvious lesions in nerve tissue that are indicative of neurotoxicity are observed occasionally. At the neurophysiological level, either diminished

responsiveness to a stimulus due to alteration of input or processing, or enhanced responsiveness to stimulation due to decreased inhibitory thresholds in the nervous system may be observed in response to exposure to a neurotoxicant. Behavioral effects of neurotoxicity can include changes in sensation; alterations in reflexes and motor function; learning, memory and mood changes; disorientation or distortions in thinking.

At each level of organization, the nervous system can be probed for information related to neurotoxicity. Evaluations can focus on a single level or combine techniques to examine two or more levels. They can be conducted immediately following an acute exposure or at some later timepoint. The effects of repeated exposure can be evaluated within the exposure period, immediately upon completion of a series of exposures or after a subsequent period without additional exposure. The results of such evaluations can be dependent on the time of measurement, with either qualitatively or quantitatively different results observed at different time points. A negative finding measured at any single timepoint may not accurately reflect the hazard potential of a compound.

Reversibility and Compensation

In agreement with the basic concepts of toxicology, irreversible effects on measures of neurotoxicity are generally considered adverse. However, when apparently reversible effects are observed, the neurotoxicologist often departs from the general principles of toxicology. Recovery of function following exposure to a neurotoxic agent and a period of evident neurotoxicity does not necessarily indicate that the neurotoxicity is reversible. Rather, it may reflect the inherent redundancy of the nervous system and result from the activation of a reserve capacity.

The relative lack of ability to regenerate tissue to replace that which has been damaged is a hallmark of the nervous system. Although recent developments indicate that this inability is not as absolute as once thought, and although work is underway on promising techniques to stimulate regeneration, nervous system damage is still considered permanent.

What then is responsible then for recovery of function? Why is rehabilitation following an insult to the nervous system successful? One answer is that the nervous system is redundant: under normal circumstances, most brain regions contain more cells than are needed to function. If some of those cells are damaged, others can begin to function in their place. This ability of some cells in the nervous system to take on a new function is referred to as plasticity.

These two features of the nervous system are necessary for survival, but they greatly complicate the process of hazard identification for the neurotoxicologist. When an observed effect following exposure to a neurotoxic agent appears to resolve, it is difficult to determine whether the effect is truly reversible. Perhaps the agent acted for a short period of time with no lasting influence. If so, the hazard classification of that agent would not be as extreme as one with a more lasting effect. Conversely, exposure to the agent may have permanently damaged cells whose function is now being performed by other cells. In the latter case, the true toxicity of the agent may be to hamper severely the organism's overall ability to respond successfully to additional challenges. Thus, the overall reserve capacity of the organism is diminished to some extent. This may not be apparent until much later and will be exceedingly difficult to associate with the initial exposure.

Delay Between Exposure and Effect

One type of delayed neurotoxicity is due to the time required for completion of underlying biochemical events. Some such effects take more time to develop than the duration of a standard hazard assessment. When the time to effect is known, observation periods are adjusted accordingly. For example, compounds related to certain organophosphates, which are known to produce delayed effects on the peripheral nervous system, are screened under protocols specifically designed to include time for the effect to develop. It is possible that many other compounds produce temporally dependent effects that are missed under standard testing protocols. In animal testing the opportunity to detect latent effects is simply not available since animals are not retained beyond the completion of a study. In humans, later clinical evaluation may document an effect, but in the absence of recent exposure information, it is difficult to link to the appropriate agent.

Another type of delayed neurotoxicity results from an exposure at one point in development that produces an effect not evident until the organism reaches another phase in its life span. Such exposures can occur in the fetus, infant, child or adult. Developmentally linked effects fall into a number of categories based upon the time of exposure or the time of the observed effect.

Certain neurotoxicities develop only if the fetus is exposed transplacentally at a specific point of development. These exposures often have no detectable effect on the mother. A classic example is methylmercury. Exposure levels that do no apparent harm to the mother can produce profound effects on the child. Some of these are apparent at birth, other effects manifest themselves only as failures at certain developmental milestones.

Exposure to neurotoxic agents during the postnatal period of rapid development can produce lasting and latent neurotoxicity. Perhaps the best documented example is provided by the relationship between early exposure to lead and later neurobehavioral function. Cognitive effects related to lead exposure in infants may not be observed until the child reaches school age. Another important developmental period during which latent neurotoxicity can be manifest is puberty. The development of secondary sex characteristics and gender appropriate sexual behavior can be influenced by gestational or postnatal exposure to neurotoxic agents. The final developmental stage is studied the least. It has been speculated, however, that exposure to neurotoxic agents compensated for over the life span may result in premature aging or rapid deterioration of function at the end of the life span (Weiss, 1990). In addition to this general acceleration of the normal aging process, a number of neurological diseases with increasing prevalence have proposed links to exposure to neurotoxic agents earlier in life. These include Parkinson's and Alzheimer's disease (Rheul, 1991).

SUMMARY

Neurotoxicity is a complex endpoint of toxicity. Over a relatively short period of time neurotoxicologists have developed a comprehensive understanding of the nervous system. With increased understanding of the structure and function of the nervous system comes enhanced ability to detect toxic effects. As the regulatory community begins to apply this knowledge increased emphasis is placed on hazard identification. Accurate hazard identification for neurotoxic compounds requires a comprehensive evaluation along many dimensions at timepoints that may be far removed from the exposure. Failure to evaluate at the appropriate post-exposure interval may result in misclassification of compounds or incorrect relative risk estimates.

REFERENCES

FCCSET (Federal Coordinating Committee on Science, Engineering, and Technology), Principles of Neurotoxicity. *Federal Register*, 1993.

OTA (US Congress, Office of Technology Assessment), *Neurotoxicity: Identifying and Controlling Poisons of the Nervous System.* OTA-BA-436, Washington DC, Government Printing Office, 1990.

Reuhl, K.R., Delayed Expression of Neurotoxicity: The Problem of Silent Damage. *Neurotoxicology*, 12:341–346, 1991.

Weiss, B., Risk Assessment: The Insidious Nature of Neurotoxicity and the Aging Brain. *Neurotoxicology*, 11:305–314, 1990.

THE NEUROTOXICITY OF CHOLINESTERASE-INHIBITING INSECTICIDES: PAST AND PRESENT EVIDENCE DEMONSTRATING PERSISTENT EFFECTS

Stephanie Padilla, Ph.D.
Cellular and Molecular Toxicology Branch, Neurotoxicology Division (MD-74B)
U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711

ABSTRACT

It is assumed that the primary mechanism of action of carbamate and organophosphate insecticides is the inhibition of an enzyme, acetylcholinesterase (AChE; EC 3.1.1.7). This enzyme normally maintains the proper level of the neurotransmitter acetylcholine in the central and peripheral nervous systems. When AChE activity is depressed due to exposure to cholinesterase-inhibiting insecticides, the ordinarily rapid breakdown of acetylcholine is retarded, causing overstimulation of target cells, which, under extreme conditions, is termed a "cholinergic crisis".

At the present time, the above indicators of acute intoxication, i.e., cholinesterase inhibition and/or clinical signs of cholinergic overstimulation, are used in the risk assessment process to regulate cholinesterase-inhibiting insecticides. There is, however, accumulating evidence both from epidemiological studies and from experimental laboratory studies that short-term exposure to some cholinesterase-inhibiting insecticides may precipitate long-term adverse effects. Recent experimental studies in our laboratory and in others have demonstrated that a single or short-term exposure to some cholinesterase-inhibiting insecticides may produce behavioral or neurochemical changes lasting for days or months, presumably outlasting the cholinesterase inhibition. This body of amassing evidence should alert those in the scientific and regulatory arenas as to the multifaceted nature of the toxicity profile of cholinesterase-inhibiting insecticides.

Organophosphate and carbamate compounds comprise the large majority of insecticides used throughout the world. These compounds are designed to inhibit an esterase, acetylcholinesterase, and presumably produce toxic responses through the short-term inhibition of this enzyme. It is commonly thought that the toxicity profile of these compounds progresses according to the following scenario: the carbamate or organophosphate inhibits acetylcholinesterase activity; because acetylcholinesterase terminates the action of acetylcholine, a neurotransmitter found in the central nervous system (CNS), neuromuscular junction, the parasympathetic nervous system, the sympathetic synapses, and the

sympathetic innervation of the adrenal and sweat glands, the concentration of acetylcholine increases in those areas, causing an overstimulation in the target cells. An ideal pesticide, of course, would not be toxic to humans and other mammals. Unfortunately, both organophosphate and carbamate compounds will cause a cholinergic crisis in humans with the following signs and symptoms: miosis, tachycardia, bradycardia, changes in blood pressure, increased sweating, lacrimation, salivation, nausea, abdominal pain, diarrhea, and muscle fasciculation. Short-term CNS effects include headache, confusion, irritability, difficulty in concentration, and emotional lability (Gershon and Shaw, 1961; Bowers et al., 1964; for a review see D'Mello, 1993). The majority of these signs and symptoms of acute poisoning usually abate as the level of cholinesterase activity returns to normal. As a consequence, organophosphate and carbamate insecticides are assumed to be relatively safe because at sub-lethal dosages they only precipitate short-term, reversible effects.

If one, however, peruses the literature on cholinesterase inhibitors as long as 30 years ago, there are indications that the above scenario cannot be presumed to be the only toxicological manifestation of cholinesterase-inhibiting insecticides. After the acute signs and symptoms have abated, there may be some residual consequences of the exposure to the cholinesterase inhibitor. For instance, it is known that a subset of the organophosphate compounds cause organophosphate-induced delayed neuropathy: a bilateral neuronal degeneration of the peripheral and central nervous system resulting in "glove and stocking" sensory and motor deficits. The present overview will not consider this neuropathy because it is assumed that the neuropathic organophosphates are identified and eliminated due to a requirement to test all cholinesterase-inhibiting insecticides for their capability to produce organophosphate-induced delayed neuropathy as part of the registration process. Even so, there are numerous reports of commonly-used insecticides producing organophosphate-induced delayed neuropathy in humans, including chlorpyrifos (Kaplan et al., 1986; Lotti et al., 1986) and methamidophos (Senanayake, 1985; McConnell et al., 1994). For a recent overview of the full spectrum (i.e., acute, intermediate, neuropathic and long-term) of postulated effects of cholinesterase-inhibiting insecticides, the reader is referred to Marrs (1993).

Many of the earlier studies on the persistent effects of cholinesterase-inhibiting compounds were summarized in a review by Karczmar (1984). Most of the long-term (and short-term) mental effects cited were due to exposure not to insecticides but to nerve agents--for example DFP (diisopropylfluorophosphate), sarin, and GB. Long-lasting effects included memory deficits, sleep disorders, hallucinations, and confusion. One of the earliest reports (Gershon and Shaw, 1961) found

that men and women exposed repeatedly to insecticides (1.5–10 years) displayed schizophrenic and depressive disorders accompanied, in some cases, by severe memory impairment and difficulty in concentration. In another early study, Metcalf and Holmes (1969) indicated that agricultural and industrial workers exposed to cholinesterase-inhibiting insecticides complained of forgetfulness, difficulty in thinking, visual problems and persistent muscular aches and pains, while no differences in other complaints (for instance, frequency of somatic disease) were noted between the exposed and control groups. Neurological examination of these workers supported the subjects' symptom reporting: the principle findings were slow thinking and calculation, memory deficits, minor coordination deficits and oculomotor imbalance. The authors concluded "It appears, therefore, that there is need for further intensive study because of the unknowns we now recognize and because of the possibility that long-term exposure to OP compounds can induce irreversible or only slowly reversible brain dysfunction." (Metcalf and Holmes, 1969). Curiously, there was no flurry of laboratory and epidemiological investigations to evaluate these allegations of long-term effects of cholinesterase-inhibiting compounds, possibly because other reports appeared to refute these observations (e.g., Durham et al., 1965).

In the early 1980s there began a resurgence of interest and experimentation regarding the long-term repercussions of acute exposures to cholinesterase-inhibiting insecticides. Midtling and coworkers (1985) found that although the number of adverse effects cited by humans subsided after poisoning, there were still effects in some workers months later when blood cholinesterase activity had returned to normal. These long-lasting effects included blurred vision, headache, weakness, or anorexia. Savage and coworkers (1988), in their comparison of acutely exposed (poisoned) and control workers, found that while a detailed neurological exam showed many endpoints were not different between the two groups, the exposed group showed significant changes in tests designed to measure abstraction, mood, and one test of motor reflexes. Differences were also apparent in neuropsychological tests of intellectual functioning, academic skills, abstraction and flexibility of thinking and simple motor skills. The authors conclude with a strong statement, but one supported by their evidence: "Results clearly indicate that there are chronic neurological sequelae to acute organophosphate poisoning," and cautioned that appropriate tests must be conducted to reveal the deficits. Moreover, this contention is supported by another study in which "poisoned" pesticide workers exhibited long-term deficits in neuropsychological performance: e.g., tests to assess memory, attention, visuomotor function, and motor skills (Rosenstock et al., 1991).

The above cited examples were for the most part studies of men and women who had been exposed acutely to high levels of cholinesterase-inhibiting compounds. Taken in concert, many investigations can be found which report adverse effects in humans, presumably due to cholinesterase inhibition, long after cholinesterase inhibition had disappeared. Can animal testing uncover the propensity of cholinesterase-inhibiting insecticides to produce long-term effects? This may be quite difficult as the majority of the long-term effects reported in the human studies were symptoms rather than signs, or endpoints which can be difficult to assess in common laboratory animals. Even in the face of these handicaps, it is noteworthy that the recently published animal literature provides some emerging indications that controlled animal studies may indeed uncover persistent, if not permanent, effects of cholinesterase-inhibiting insecticides.

Using screening tests recommended by the U.S. Environmental Protection Agency (Functional Observational Battery and motor activity), Moser found that the effects of carbamate insecticides lasted for days after cholinesterase levels have returned to normal (Moser, 1994; Moser, Padilla, and Nostrandt, unpublished data). Using the same screening battery, Ehrich's group noted that adverse effects of both organophosphate and carbamate insecticides were still detectable weeks after an acute dose (Ehrich et al., 1993). Interestingly, in some cases the profile of these longer-term clinical and/or behavioral alterations differed from the short-term effects, similar to what was often reported in the human literature.

Moreover, there are two reports which would indicate that the neurochemistry of animals exposed to either chlorpyrifos (Pope et al., 1992) or fenthion (Tandon et al., 1994a;c) may be permanently altered after one exposure. Months after a single sc dose of chlorpyrifos, when both muscarinic receptor density and cholinesterase activity had returned to normal levels, rats displayed enhanced hyperactivity when challenged with a muscarinic antagonist, scopolamine (Pope et al., 1992). Scopolamine treatment normally elicits some degree of hyperactivity in rats, but the rats previously treated with chlorpyrifos showed an exaggerated response to this pharmacological challenge. In a similar vein, numerous studies have showed that after a single dose of fenthion intracellular communication by retinal muscarinic receptors appeared to be permanently altered (56 days after an acute dose or 104 days after cessation of a repeated dosing regimen, after cholinesterase activity and muscarinic receptor density had returned to normal) (Tandon et al., 1994a;b;c).

In conclusion, there seems to be firm evidence in the human literature, which is now supported by an emerging experimental animal literature, of long-term effects of cholinesterase-inhibiting

insecticides. What is not clear is whether a substantial amount of cholinesterase inhibition is a prerequisite for the precipitation of these persistent effects, as the great majority of the human subjects had experienced a "poisoning" episode and animals were usually given dosages which produced more than 50% inhibition of brain cholinesterase activity. Also important among the list of other unanswered questions is whether long-term, low-level exposure would also produce these effects (see Davies, 1990) and whether these effects would be more severe in developing animals where both acetyl-and butyrylcholinesterase may play an integral role in the development of the nervous system (e.g., Layer, 1991).

REFERENCES

Bowers, M.B., Goodman, E., and Sim, V.M. 1964. Some behavioral changes in man following anticholinesterase administration. *J. Nerv. Ment. Dis.* 138:383–389.

Davis, J.E. 1990. Neurotoxic concern of human pesticide exposures. *Amer. J. Indust. Med.* 18:327–331.

D'Mello, G.D. 1993. Behavioural toxicity of anticholinesterases in humans and animals-A review. *Human and Experimental Toxicology* 12:3–7.

Durham, W.F., Wolfe, H.R., and Quinby, G.E. 1965. Organophosphorus insecticides and mental alertness. *Arch. Environ. Health* 10:55–56.

Ehrich, M., Shell, L., Rozum, M., and Jortner, B.S. 1993. Short-term clinical and neuropathologic effects of cholinesterase inhibitors in rats. *J. Amer. Coll. Toxicol.* 12:55–67.

Gershon, S., and Shaw, F.H. 1961. Psychiatric sequele of chronic exposure to organophosphorus insecticides. *The Lancet* 1:1371–1374.

Kaplan, J.G., Kessler, J., Pack, D.R. and Schaumburg, H.H. 1986. Dursban causes peripheral neuropathy. *Neurology* 36:176.

Karczmar, A.G. 1984. Acute and long lasting central actions of organophosphorus agents. *Fund. Appl. Toxicol.* 4:S1–S17.

Layer, P.G. 1991. Cholinesterases during development of the avian nervous system. *Cell. Mol. Neurobiol.* 11:7–33.

Lotti, M., Moretto, A., Zoppellari, R., Dainese, R., Rizzuto, N., and Barusco, G. 1986. Inhibition of lymphocytic neuropathy target esterase predicts the development of organophosphate-induced delayed polyneuropathy. *Arch. Toxicol.* 59:176–179.

McConnell, R., Keifer, M., and Rosenstock, L. 1994. Elevated quantitative vibrotactile threshold among workers previously poisoned with methamidophos and other organophosphate pesticides. *Amer. J. Indust. Med.* 25:325–334.

Marrs, T.C. 1993. Organophosphate poisoning. *Pharmac. Ther.* 58:51–66.

Midtling, J.E., Barnett, P.G., Coye, M.J., Velasco, A.R., Romero, P., Clements, C.L., O'Malley, M.A., Tobin, M.W., Rose, T.G., and Monosson, I.H. 1985. Clinical management of field worker organophosphate poisoning. *West. J. Med.* 142:514–518.

Metcalf, D.R., and Holmes, J.H. 1969. EEG, psychological, and neurological alterations in humans with organophosphorus exposure. *Ann. N.Y. Acad. Sci.* 160:357–365.

Moser, V.C. 1994. Comparisons of the acute effects of cholinesterase inhibitors using a neurobehavioral screening battery in rats. *The Toxicologist* 14:241.

Pope, C.N., Chakraborti, T.K., Chapman, M.L. and Farrar, J.D. 1992. Long-term neurochemical and behavioral effects induced by acute chlorpyrifos treatment. *Pharmacol. Biochem. Behav.* 42:251–256.

Rosenstock, L., Keifer, M., Daniell, W.E., McConnell, R., and Claypoole, K. 1991. Chronic central nervous system effects of acute organophosphate pesticide intoxication. *The Lancet* 338:223–227.

Savage, E.P., Keefe, T.J., Mounce, L.M., Heaton, R.K., Lewis, J.A., and Burcar, P.J. 1988. Chronic neurological sequelae of acute organophosphate pesticide poisoning. *Arch. Environ. Health* 43:38–45.

Senanayake, N. 1985. Polyneuropathy following insecticide poisoning. *J. Neurol.* 232:203.

Tandon, P., Padilla, S., Barone, Jr., S., Pope, C.N., and Tilson, H.A. 1994a. Fenthion produces a persistent decrease in muscarinic receptor function in the adult rat retina. *Toxicol. Appl Pharmacol.* 125:271–280.

Tandon, P., Pope, C.N., Barone, S., Jr., Boyes, W., Tilson, H.A., and Padilla, S. 1994b. Fenthion treatment produces tissue-, dose- and time-dependent decreases in muscarinic second messenger response in the adult rat CNS. To be presented at the 5th International Meeting on Cholinesterases in Madras, India, Sept., 1994.

Tandon, P., Willig, S., Pope, C.N., Padilla, S., and Tilson, H.A. 1994c. Dose-response study of the tissue-specific effects of fenthion on receptor function in rat CNS. *The Toxicologist* 14:257.

ACKNOWLEDGMENTS

The article has been reviewed by the Health Effects Research Laboratory, U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Agency, nor does mention of trade names and commercial products constitute endorsement or recommendation for use.

SESSION III

EFFECTS OF EXPOSURE PATTERNS OF DOSE/RESPONSE

What do we mean by . . . DOSE?

Melvin E. Andersen, Ph.D.

ICF Engineers, KS Crump Division
1 Copley Parkway Suite 102, Morrisville, NC 27560

ABSTRACT

Accurate risk assessments for chemical health hazards are closely tied to the ability to predict the conditions under which toxic effects are likely to occur in exposed people based on observations of toxicity in experimental animals. These predictions require (1) knowledge of the dose–response characteristics for the effect and (2) the ability, based on mechanistic considerations, to extrapolate to expected incidences at much lower doses in exposed people. An important concept for these extrapolations is the definition of 'dose', i.e., to what measure of 'dose' should we relate toxicity and how does the chosen measure of dose extrapolate across species? This brief commentary looks at the evolution of our understanding of both the chemical and biological aspects of the concept of tissue dose.

INTRODUCTION

Over the past several decades, the concept of dose as applied to toxicological studies has changed considerably (Andersen, 1987). Initially, 'dose' simply meant the concentration in the atmosphere in inhalation studies times the duration of the study or the amount ingested or instilled into the gastrointestinal tract in oral dosing studies. Both these quantities are simply measures of delivered dose as mg instilled or mg inhaled. Empirical correlations were then used to evaluate the relationship of dose and response with little appreciation of the detailed biological interactions of the test compounds (Figure III-1). Many chemicals, however, are metabolized to toxic compounds by saturable, enzymatic processes. With some of these chemicals the relationships between administered amounts of compound and response were complex (Gehring et al., 1977; Andersen, 1981). These complex dose–response curves were partially resolved by relating response to metabolized dose or to concentrations of metabolites reached in target tissues. This first expansion of our definition of dose, from administered to target tissue dose, led to a much increased appreciation of the role of pharmacokinetic processes in regulating toxic chemical delivery to target tissues and in subsequent toxicity (Andersen, 1981). In

recent years the concept of dose has expanded once again to focus increasingly on interspecies differences in the biological processes involved in cell constituent biosynthesis, macromolecule turnover, and homeostatic control mechanisms (Figure III-2). In this paper Figure III-2 serves as a focal point of our attempts to refine the definition of "dose", especially in reference to interspecies extrapolation of expected risks.

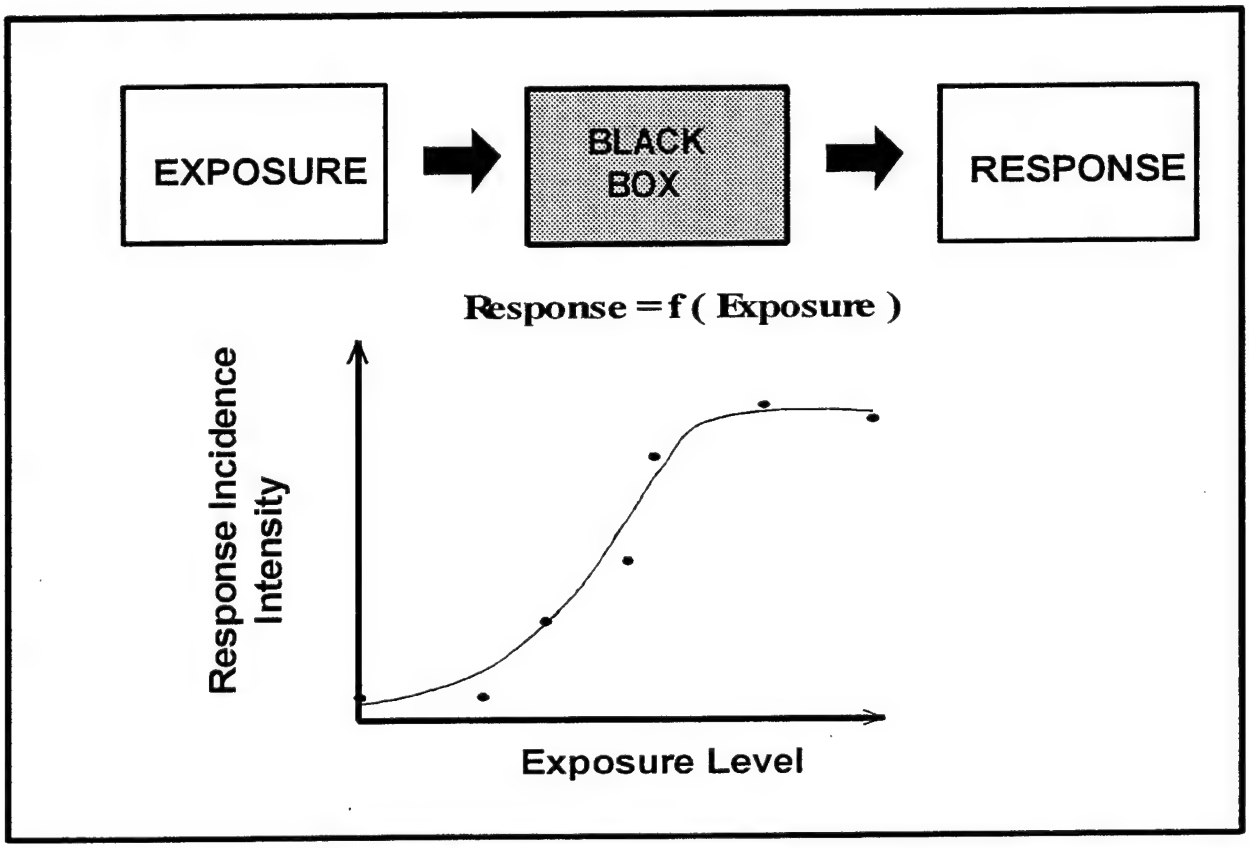


Figure III-1. Developing Empirical Dose–Response Curves for Effects of Chemicals.

Originally, toxic responses were simply related to an external measure of delivered dose with little understanding of the complex series of steps leading from exposure, to accumulation of chemical in tissue, to interactions with tissue components, and eventually to both toxic and accommodative cellular responses. The body in essence was treated as if it were a black-box.

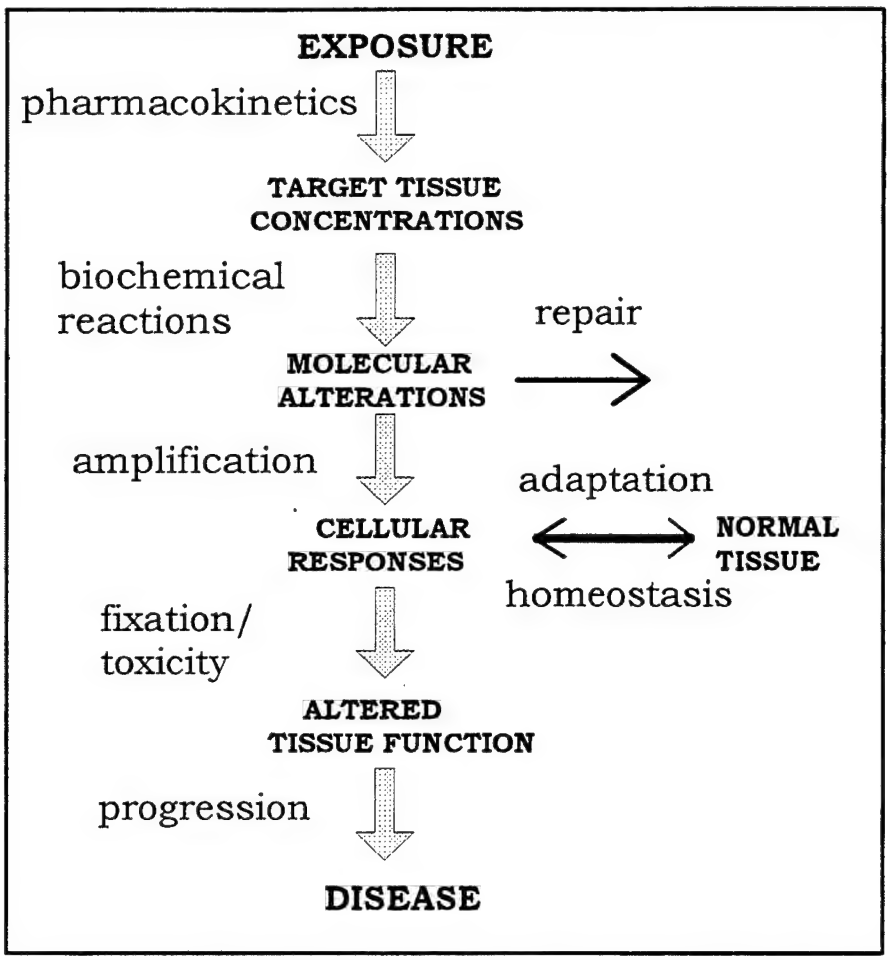


Figure III-2. The Chemical and Biological Processes Involved in the Expression of Toxic Responses.

The overall dose–response curve for a particular effect is determined by the interplay of a series of physico-chemical and biological processes. In evaluations in a single species the role of the biological processes have frequently been ignored and improvements in dose–response behaviors obtained by explicitly considering dose-to-tissues and the binding or interaction of chemical with tissue macromolecules. Interspecies extrapolation of dose–response behaviors is greatly improved when knowledge is available regarding the variation of both the chemical and the purely biological processes in various animal species.

MEASURES OF DOSE

Administered Dose–Achieved Tissue Concentrations

In evaluating drug efficacy, the prevalence of response or intensity of a particular outcome is frequently plotted versus the administered dose as in Figure III-1. The response of an individual to alcohol, for instance, is frequently specified in relation to the number of drinks required to reach a particular state of inebriety. This is a measure of administered dose, as gm alcohol/kg body weight. This measure of dose does not take the time dimension, the dose rate, into effect. Three drinks in a 50-kg woman taken over 45 minutes may give a blood alcohol of 0.1%, the legally defined state of intoxication; three drinks spaced at 12-hour intervals will not. In this case, the concentration of ethanol in blood approximates the target site concentration in the nervous system and referencing effects to concentration eliminates the obvious confounding of response by alterations in dose-rate. With CNS depressants like alcohol or volatile anesthetics, the achieved concentration in the target tissues is a more important correlate of toxicity than is administered dose. Figure III-3 indicates some of the diverse measures of dose that might be appropriate for varying chemicals and toxic endpoints.

Achieved Concentrations-Metabolized Dose

Metabolism often terminates drug action, as it does with ethanol. Saturation of metabolism may lead to unanticipated accumulation of drug and then to toxicity due to enhanced tissue concentrations or prolonged exposure of target tissues to active xenobiotics (a measure of cumulative tissue dose). With drugs, therapeutic regimens are generally designed to maintain first-order elimination behavior of efficacious compounds to avoid these dosing problems. In contrast to drugs, many toxic compounds are not directly effective. They are metabolized via saturable enzymatic processes to toxic compounds which interact and interfere with cellular processes. Vinyl chloride is metabolized to a DNA alkylating epoxide, leading eventually to hemangiosarcoma of the liver in rats, mice, and occupationally exposed people. The appropriate dose-measure with vinyl chloride is metabolized dose (mg/kg). Because metabolism is saturable, the response reaches a plateau and does not increase with increasing dose (Maltoni et al., 1974; Gehring et al., 1978). When toxicity or carcinogenicity is related to mg/kg metabolized, it is important to add the caveat that the correlation holds for one particular animal species, in one particular exposure situation. This situation is reminiscent of the confounding of alcohol intoxication by dose-rate and is partially clarified by specifying the mg metabolized/day and the duration of the experiment in days, for instance. For broader extrapolation, more information is required on adduct repair rates and other biological processes occurring over time.

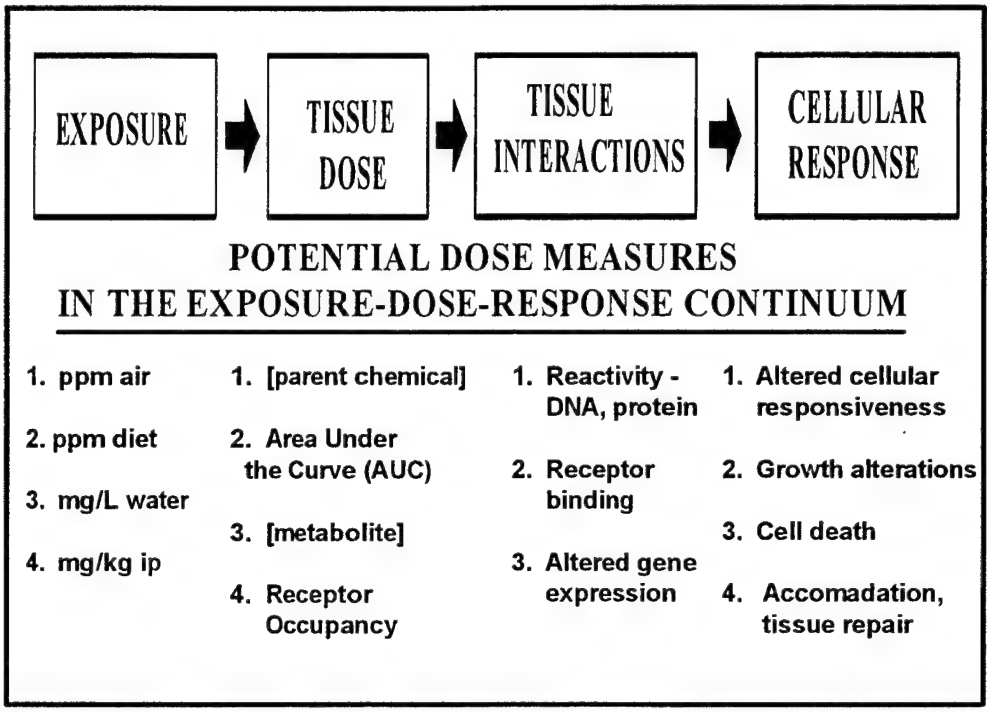


Figure III-3. Measures of 'Dose' that Include Biological and Chemical Characteristics of Chemical Intoxication, Biological Response, and Adaption Will Be More Appropriate When the Goal of the Toxicity Studies Include Interspecies Extrapolation.

Metabolized Dose-Cellular Protective Mechanisms

The concentration of biological substances within cells can also be a critical determinant of the dose–response behavior. Epoxide intermediates react with glutathione — GSH (Watanabe et al., 1976). A more correct measure of dose for vinyl chloride, discussed above, might be the amount of epoxide (mg/kg) escaping glutathione conjugation that is available for tissue reactions. Vinylidene chloride (VDC: 1,1-dichloroethylene) is metabolized to an extremely hepatotoxic intermediate, presumably chloroacetylchloride produced by rearrangement of an epoxide intermediate. A physiologically—based pharmacokinetic (PB—PK) model was developed for VDC to evaluate the role of GSH in acute VDC toxicity and described the unusual dose-dependence of VDC hepatic toxicity (D'Souza et al., 1987). The model described the amount of metabolite formed that escaped GSH conjugation. Two critical determinants of the protection by glutathione were purely biological in origin, the basal level of hepatic glutathione and the resynthesis rate of the co-factor. In this case, the dose measure, i.e., the amount of reactive metabolite available for tissue interactions, is intimately affected by a biological process, the GSH resynthesis rate. The inclusion of the glutathione synthesis rate is

especially important in extrapolating from one species to another and in dealing with expected damage for various exposures at various dose rates. For a given amount of metabolized epoxide, production at a slower rate may be without consequence if GSH resynthesis is rapid enough to maintain adequate GSH stores.

Tissue Exposure to Metabolites-Tissue Repair Rates

2,5-Hexanedione, a stable metabolite of both *n*-hexane and methyl-*n*-butylketone, is neurotoxic, causing a central, peripheral, distal, neuropathy. The proposed mechanism involves the reaction of 2,5-HD with neurofilament protein (NFP), cross-linking of the NFP, and impairment of neurofilimentary transport processes. The measure of dose for this adverse effect is expected to be the integrated exposure, area under the curve (AUC), for 2,5-HD in the nervous tissues. The development of toxicity depends on exposure of the NFP to 2,5-HD, NFP cross-linking, repair of NFP cross-links, repair and replacement of NFP themselves within the cell, and, for the case of overt response, the repair and recruitment of nerve cells to restore normal function. The adduction rates of 2,5-HD with NFP are more chemical specific; the repair, regeneration, and recruitment processes are more biological in origin. The latter processes are important in any species in determining the overall dose response behavior for 2,5-HD neurotoxicity. However, the role of these biological processes in determining the dose-response behavior tends to be ignored in studies in a single species of test animal. As the focus shifts from the single species to the issue of extrapolation to people based on the animal results, these processes have to be explicitly defined and their role in giving rise to the 'dose–response' behavior more extensively investigated.

DOSE AS EFFECTED BY BIOLOGICAL PROCESSES

Cell Replication and Repair

Toxicity with VDC is eventually related to the loss of viable hepatocytes. Cell number is determined by the initial number of hepatocytes, the rate of VDC-induced cell death and, the rate of replenishment of new cells by cell replication With overtly toxic exposures, the birth rate of new cells may be a critical determinant in the outcome, death or survival of an exposed individual, in a population of animals. Mehendale et al. (1990) have shown that Kepone inhibits cell-replication after hepatotoxic exposures and increases the lethality of CCl4 and other hepatotoxic compounds. In extrapolating across species, the recovery of function may be exquisitely dependent on these fundamental biological processes of cell replication in various species and tissues. Toxicity, a balance between injury and recovery, is dependent on all of these processes.

Homeostasis-A Chemical-Biological Standoff

Chemicals may affect living systems by reacting with covalently with tissue constituents or by interacting non-covalently with receptors to alter cellular function. However, living organisms are highly adaptive. Initial insults usually lead to normal, intensified responses at the organism level to maintain normal function. This process is referred to as homeostasis: keeping the status very nearly quo even though there are alterations in inputs in the system. A good example is found with the hepatocarcinogenicity of phenobarbital (PB). PB is an hepatic mitogen and causes a brief period of cell proliferation in the first weeks after the initiation of dietary PB exposure. However, cell proliferation cannot continue indefinitely. Following the brief surge of proliferation, there is an increased elaboration of an inhibitory growth factor, TGF-β1 that inhibits further hepatocyte proliferation, even though the chemical signal, in the form of PB, is still present. These homeostatic responses are necessary components of the well-regulated biological machinery of the body. With most compounds, sufficiently high concentrations may overwhelm the homeostatic control and eventually lead to overt toxicity. For this reason, the slopes of the dose response curves may vary as with any control process. At low concentrations the new input leads to little change in the system output (a low gain). When homeostasis begins to fail, the slope increases with larger changes in unit output for alterations in the input signal. This general behavior should apply, for instance, to chemicals which act as exogenous hormones by binding natural receptors and eliciting a specific biological response based on the concentration of the receptor-ligand complex. Examples here may include dioxin and the Ah receptor (Mills et al., 1993). In any case, the linearity of the complex biological response (cell division or differentiation) may be quite different in areas of dose where homeostatic mechanisms are adequate and change dramatically at higher doses.

Homeostasis and Injury

Homeostatic mechanisms may also establish altered conditions in the body that enhance the possibility of disease. Jirtle and co-workers (1991) have proposed a model with PB to explain its hepatic carcinogenicity. PB is not genotoxic and appears to function as a promoter. Jirtle proposed that the increased concentrations of inhibitory growth factor TGF-β1 creates a negative selection environment and encourages the outgrowth of spontaneously initiated cells that are TGF-β1 resistant. Thus, with PB the homeostatic response appears to create an environment conducive to growth of intermediate cell lesions that increase the risk for conversion of the intermediate cells to carcinomas. The lesson with PB relates to the interplay of chemical concentrations in tissue, physical-chemical

interactions, and biological processes, including accommodation and homeostatic controls, that are involved in the overall response and in the dose–response behavior for the observed response. The interplay of all these factors with PB leads to a complex U-shaped dose response curve for tumor promotion (i.e., creation of enzyme altered foci in rat livers).

CONCLUSIONS

The appropriate dose measure in toxicity studies will vary depending on the chemical, its biological interactions with the host organism, and the biological responses normally involved in cellular and tissue homeostasis in the animal (Figure III-3). In dose–response extrapolation the interspecies dependence of both pharmacokinetic and pharmacodynamic processes have to be carefully considered in estimating the expected responses in people. It is imperative that the richness of the concept of "dose" be kept in mind in our toxicity studies to avoid narrow, glib, and potentially misleading definitions that focus simply on some property of the test compound and its direct interactions with cellular components.

Today, we see more clearly than ever before that differences in dose–response relationships across species are as much determined by biological differences in repair rates of lesions and homeostatic responses as they are by delivery of toxic chemicals to target tissues. Especially, when discussing equivalent dose across species both pharmacokinetic and pharmacodynamic processes become important. We are constantly challenged to define dose in biologically meaningful ways in order to assist in low dose and interspecies extrapolations in risk assessment-oriented research. The evolution of an integrated chemical/biological concept of dose will continue as the sophistication in examining toxicological problems is enhanced by improved mechanistic understanding of a variety of toxic phenomena at the molecular, cellular, organism, and population level.

REFERENCES

Andersen, M.E. (1981). Saturable Metabolism and Its Relationship to Toxicity. Critical Reviews in *Toxicology* **9**:105–150.

Andersen, M.E. (1987). Tissue Dosimetry in Risk Assessment: What's the Problem Here Any Way? *Drinking Water and Health* **8**:8–23.

Andersen, M. E., Mills, J.J., Jirtle, R.L., and Greenlee, W.F. (1994). Negative selection in hepatic tumor promotion in relation to cancer risk assessment. *Toxicology, in press.*

Conolly, R.B. and Andersen, M.E. (1991). Biologically-Based Pharmacodynamic Models: Tools for Toxicological Research and Risk Assessment. *Ann. Rev. Pharmacol. Toxicol.*, **31**:503–523.

D'Souza, R.W. and Andersen, M.E. (1988). Physiologically-based Pharmacokinetic Model for Vinylidene Chloride. *Toxicol. Appl. Pharmacol.* **95**:230–240.

Gehring, P.J., Watanabe, P.G., and Park, C.N. (1978). Resolution of dose–response toxicity data for chemicals requiring metabolic activation: Example— vinyl chloride. *Toxicol. Appl. Pharmacol.* **44**:581–591.

Gehring, P.J., Watanabe, P.G., and Young, J.D. (1977). The relevance of dose-dependent pharmacokinetics in the assessment of carcinogenic hazard of chemicals. In: *Origins of Human Cancer, Book A: Incidence of Cancer in Humans* (Cold Spring Harbo Conferences on Cell Proliferation, v. 4]. H.H. Hiatt, J.D. Watson, and J.A. Winsten, eds. Cold Spring Harbor Laboratory, Cold Spring- Harbor, New York, 187–203.

Jirtle R.L., Meyer S.A., and Brockenbrough J.S. (1991). Liver tumor promoter phenobarbital: A biphasic modulator of hepatocyte proliferation. In: *"Chemically induced cell proliferation: implications for risk assessment."* Wiley-Liss Inc., 209–216.

Maltoni, C. (1975). The value of predictive experimental bioassays in occupational and environmental carcinogeneses. An example: vinyl chloride. Ambio, **4**:18–23.

Mehendale, H. (1987). Potentiation of halomethane hepatotoxicity by chlordecone: a hypothesis for the mecahanism. **33**:289–299.

Mills, J.J. and Andersen, M.E. (1993). Dioxin Hepatic Carcinogensis: Biologically Motivated modeling and Risk Assessment. *Toxicol. Ltrs.*, **68**:177–189.

Watanabe, P.H., McGowan, G.R., Madrid, E.O., and Gehring, P.J. (1976). Fate of ^{14}C-vinyl chloride following inhalation exposure in rats. *Toxicol. Appl. Pharmacol.* **37**:49–59.

INCORPORATION OF TEMPORAL FACTORS INTO PHYSIOLOGICALLY-BASED KINETIC MODELS FOR RISK ASSESSMENT

Ellen J. O'Flaherty, Jaroslav Polák, and Michelle D. Andriot
Department of Environmental Health, University of Cincinnati College of Medicine
3223 Eden Avenue, Cincinnati, Ohio 45267-0056

INTRODUCTION

Physiologically-based models offer a unique opportunity to incorporate time dependence into kinetic simulations. They are based on physically measurable physiologic functions and anatomic characteristics. The dependence of many of these parameters on age and/or body size is already known, and the age-dependence of many chemical absorption and disposition characteristics — for example, fractional absorption and renal clearance — can also be determined. These relationships can be incorporated directly into a physiologically-based model to provide a rational foundation for the prediction of age-dependent kinetic behavior.

METHOD

Description of time-dependent or age-dependent changes in key parameters of a physiologically-based model requires, first, identification of the parameters to whose values the behavior of the model is most sensitive; second, measurements that document the changes in those parameters with age or over time; and third, translation of the temporal behavior of the discrete data points into a mathematical description of a smooth curve. The first requirement can be met by sensitivity analysis, but with a simple model it can often be adequately achieved manually by working interactively with the model. If the model accurately reflects the underlying anatomy and physiology, model sensitivity analysis is equivalent to identification of those anatomic and physiologic parameters that control the kinetic behavior of the chemical. If, however, the model is faulty in some key respect, model sensitivity analysis can be misleading. In some cases, as for example during pregnancy, the characteristics of an altered physiologic state have already been well defined both qualitatively and quantitatively, and the simplest (if not the most parsimonious) approach may be to include all the temporal changes that are understood to characterize the particular physiologic state. An advantage of this method is that the resulting physiologically-based model is more likely to be generalizable from one chemical to another, since key anatomic and physiologic parameters may vary from one chemical to the next.

Data for a surprising number of age- and time-dependencies can be found in the literature. Sometimes, as for basic functions like glomerular filtration rate or respiratory rate, dependence on age is well-characterized for the entire period from birth to adulthood (Figure III-4). Often, however, there are gaps in these data sets. Data may have been collected for infants and adults but not for children (Table III-1). The rapid shifts taking place in certain functions during adolescence may be poorly characterized both as to their quantitative behavior and as to the independent variables that control that behavior.

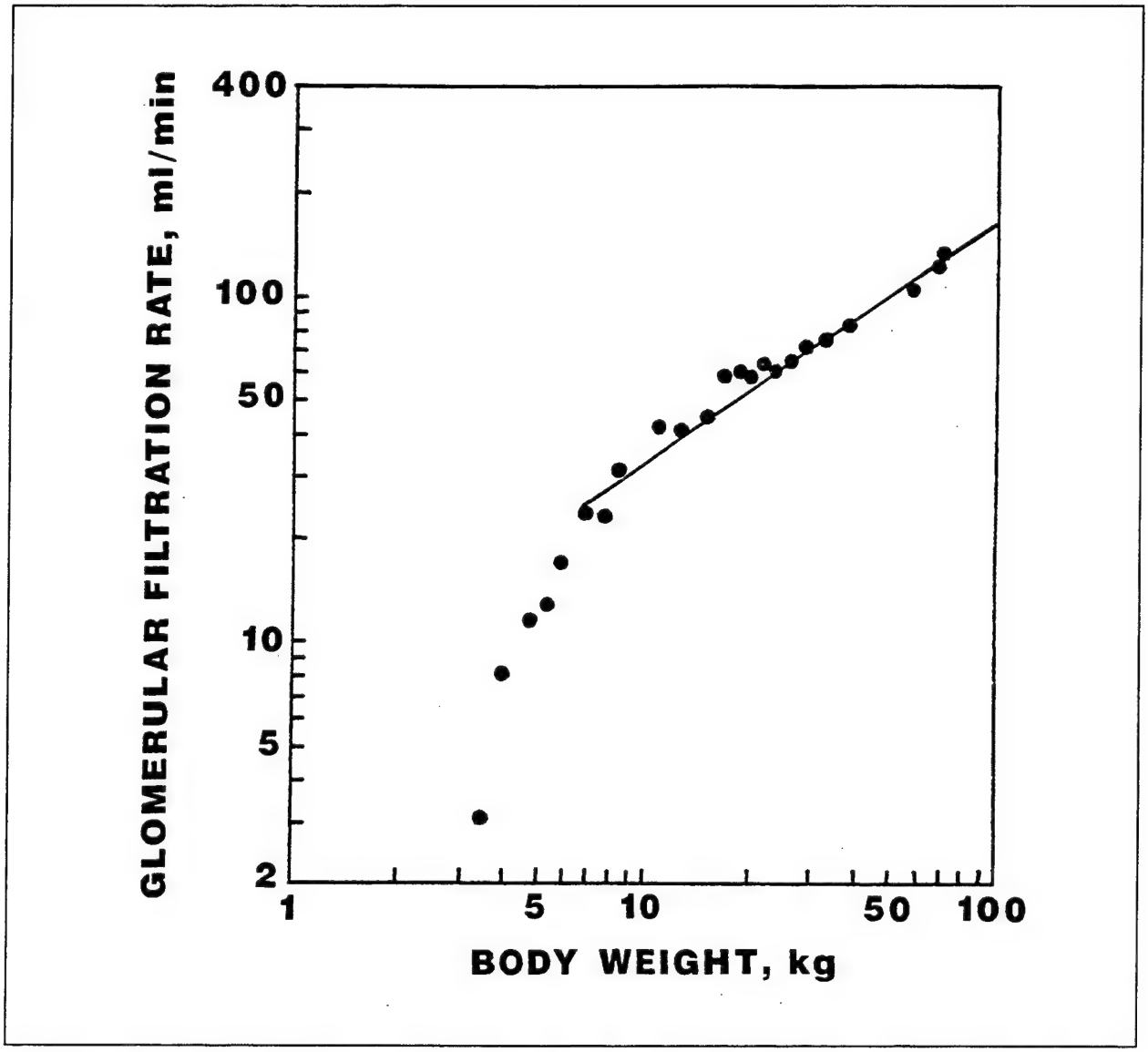


Figure III-4. Dependence of Glomerular Filtration Rate on Age. Data from Johnson et al., 1987. The line shown is an allometric fit to glomerular filtration rates for body weights above 10 kg.

Table III-1. Fractional Lead Absorption by Humans from a Normal Diet.

Number of Subjects	Age	Percent Absorption	Reference
12	5 mo–2 yr	42	Ziegler et al., 1978
8	3 mo–8.5 yr	53	Alexander, 1974
4	Adult	10.3 ± 2.2	Rabinowitz et al., 1980
8	Adult	10.0 ± 2.5	Watson et al., 1986

Translation of temporal behavior into a mathematical expression is usually not difficult. Figure III-5 illustrates the shapes of a number of common mathematical expressions that lend themselves to biological applications. The exponential function suggests the growth of the chorioallantoic placenta during gestation in the mouse. A polynomial expression describes well the increase in size of certain maternal tissues in either human or rodent pregnancy. Rodent growth (body weight) follows a logistic curve. The sine function is useful for biphasic behaviors, like the increase in maternal cardiac output during early pregnancy that is followed by a slight decrease towards the end of pregnancy. Functions may be combined to fit more complex biologic curves. An example, of the fit of a combined hyperbola (for growth during early childhood) and logistic function (for growth into adulthood) to the growth curve for Standard Man is shown in Figure III-6.

The results are given of two applications in which physiologically-based models incorporating temporal dependence of model parameters were used to simulate changes in tissue concentrations. Both scenarios involve exposure to lead.

RESULTS

Observations Made During Rapid Growth. In many studies of the bioavailability of lead from soil ingested by children, lead salts or lead-containing soils are admixed with the diet of laboratory animals, and blood and tissue lead concentrations are determined after an appropriate exposure period. Usually, the animals used in these studies are rapidly-growing weanlings or juveniles. When growth is incorporated into a physiologically-based model of lead metabolism, the model can be used to assist in interpretation of results from studies with growing animals.

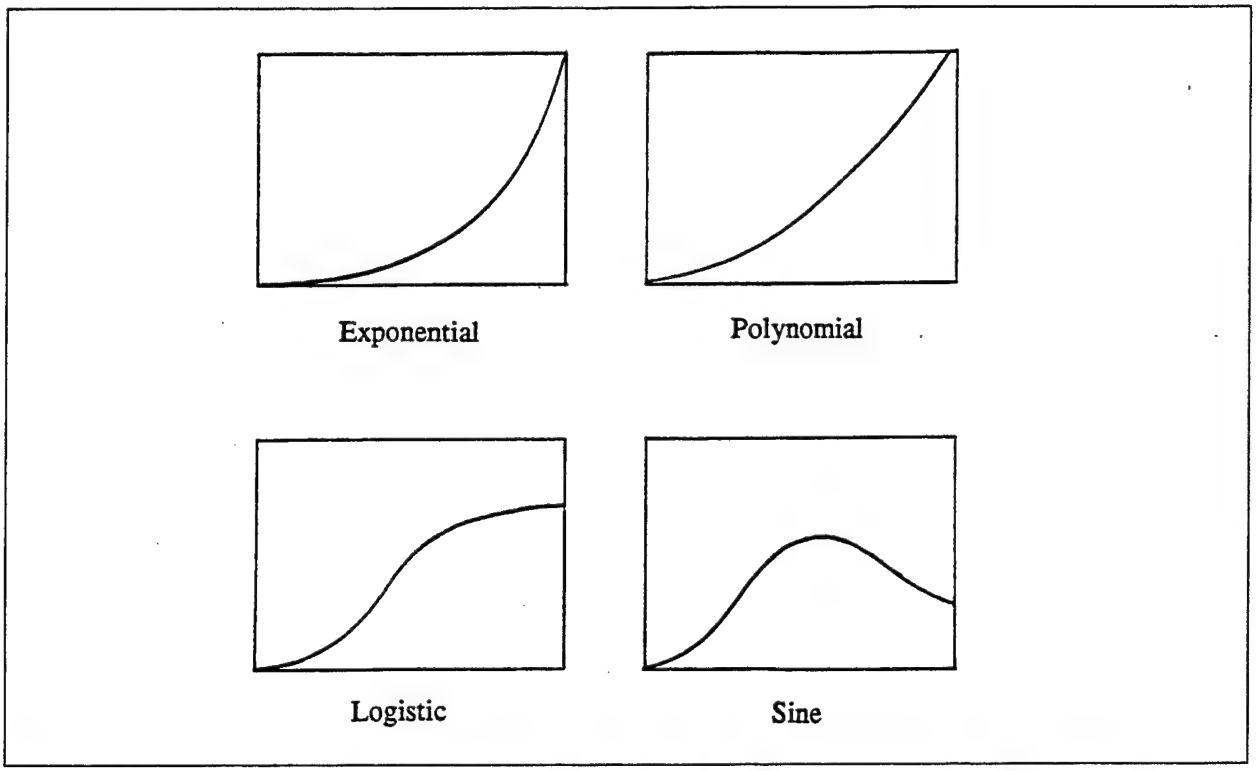


Figure III-5. Examples of Mathematical Functions with Biologically Useful Forms.

A physiologically-based model of lead metabolism in the growing and mature rat has been developed and validated (O'Flaherty, 1991). The model was applied to concentration data from rat feeding studies (Freeman et al., 1992). Young male and female Sprague-Dawley rats, 7–8 weeks old at the start of the study, were given one of two similar mining waste-contaminated soils, mixed in a purified diet at a total of 8 different dose levels, for 30 consecutive days. Growth, which was rapid during the study, was not altered by lead exposure. Lead concentrations in blood, liver, and bone were assayed at the end of the 30-day exposure period. The physiologically-based model incorporating growth was fit simultaneously to the concentration data for blood and bone by optimizing fractional bioavailability but leaving all other parameters unchanged. The results are shown in Figure III-7. The two lowest dietary lead levels resulted in tissue lead concentrations indistinguishable from background concentrations. At higher dietary lead levels, absolute bioavailability (which represents bioaccessibility in the gastrointestinal tract coupled with fractional absorption) ranged from 1.6% downward to 0.3% in the male rats and from 0.90% downward to 0.24% in the female rats, systematically decreasing as exposure increased. At all concentrations tested, the females absorbed a smaller percentage of the lead intake than the males did. When lead intake into the gastrointestinal tract was expressed in terms of

mg/kg/day, however (data not shown) fractional uptake at a given dose rate was not demonstrably different between males and females. By application of the model, which takes into account the rapid growth of the rats during the study period, multiple lead concentration measurements are synthesized into an internally consistent picture of the dependence of lead bioavailability on gender and on concentration in the gastrointestinal tract.

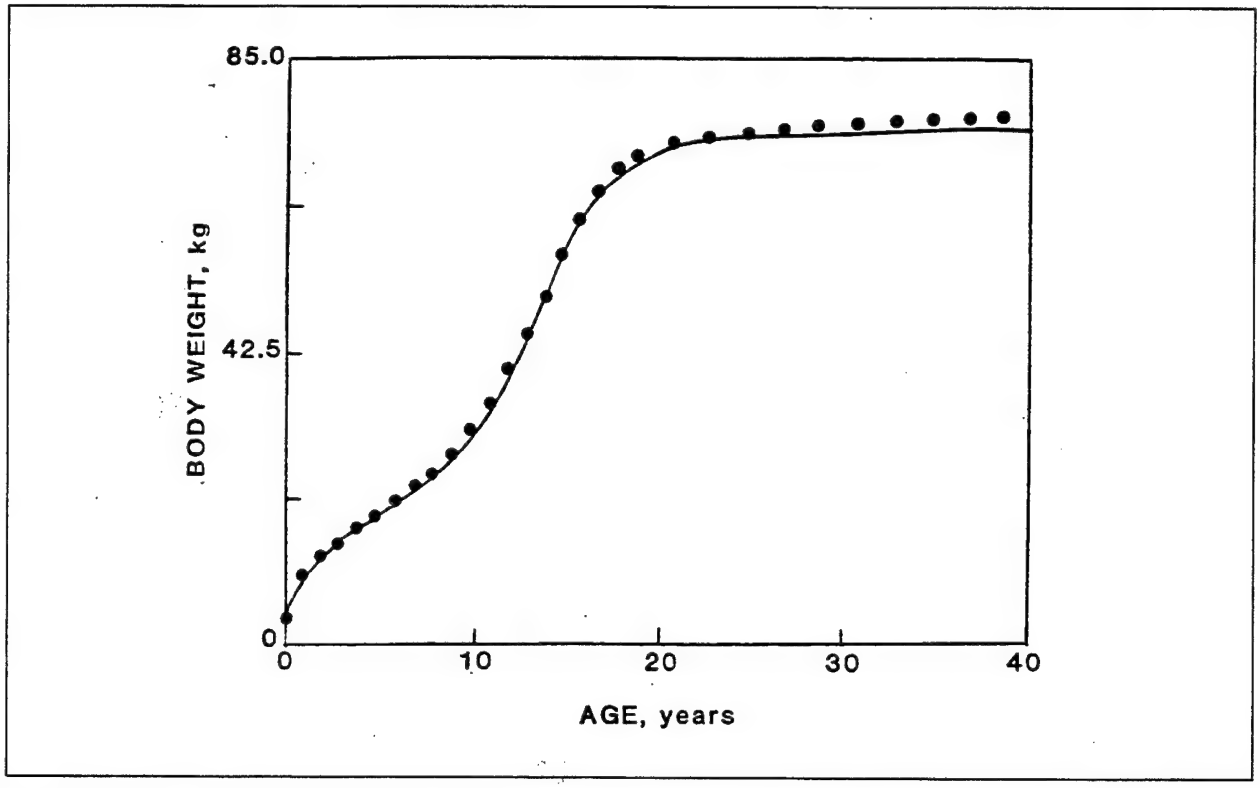


Figure III-6. The Growth Curve for Standard Man (solid points) as Fit by the Sum of a Hyperbola and a Logistic Function.

Observations Made During Pregnancy. Pregnancy is associated with marked changes in many tissue volumes and blood flows. Maternal tissues (uterus, mammary glands, fat) increase in mass, and total blood flows to these tissues increase accordingly. Plasma volume increases and then tends to decrease late in pregnancy as fluid shifts from plasma to extracellular space. Cardiac output increases and then decreases. Gastrointestinal motility slows. These and other changes during human pregnancy have been well documented, perhaps most notably in the work of Hytten and Leitch (1971). Figure III-8 illustrates the changes in hematocrit that result from shifts in plasma and red cell volumes during pregnancy. Another change that may occur is an increase in bone resorption in late pregnancy in order to maintain stable maternal levels of plasma calcium for utilization by the developing fetus.

Concern has been expressed that increased bone resorption in previously lead-exposed women could release biologically significant amounts of lead into maternal plasma at a time when the fetus in particularly sensitive to the developmental effects of lead and also most able to accumulate persistent stores of lead in newly-mineralizing bone (Pounds et al., 1991).

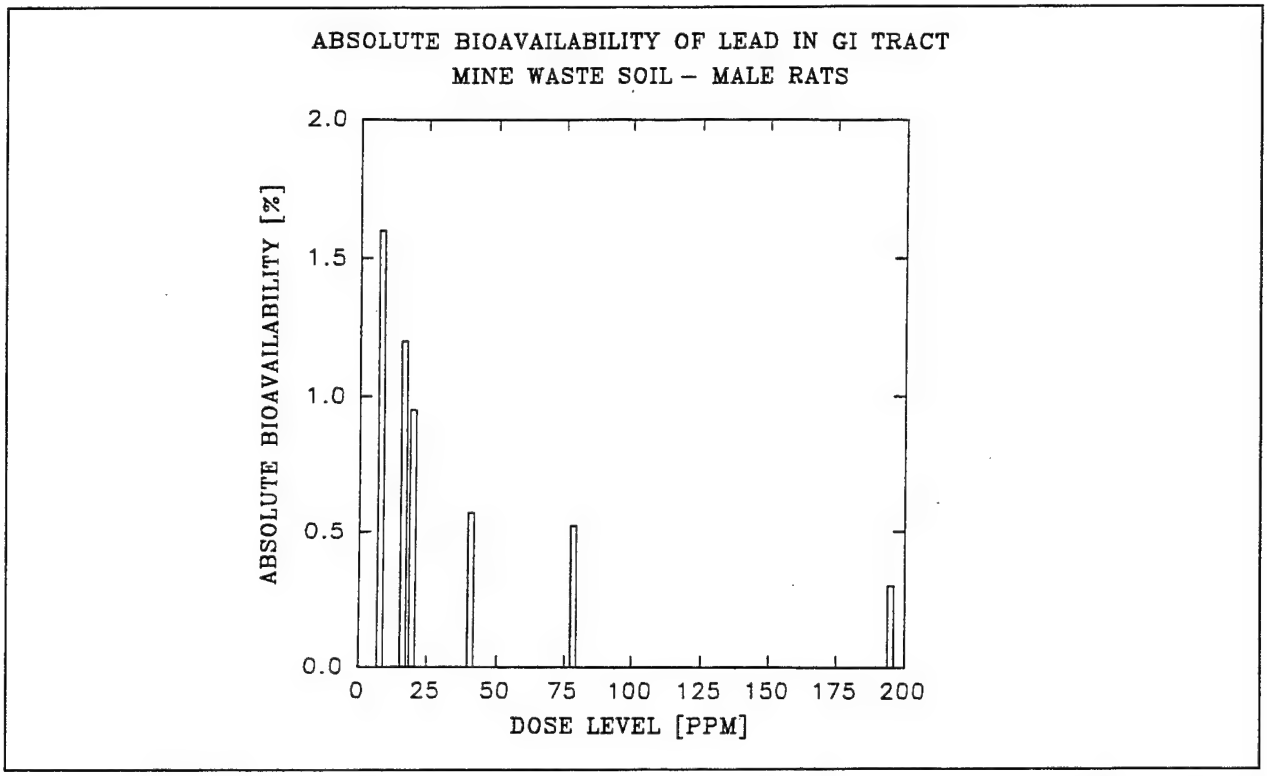


Figure III-7. Fractional Bioavailability of Lead From Mining Waste Soils Added to Diet of Male Rats. See text for Details of study. Dose levels from Freeman et al., 1992.

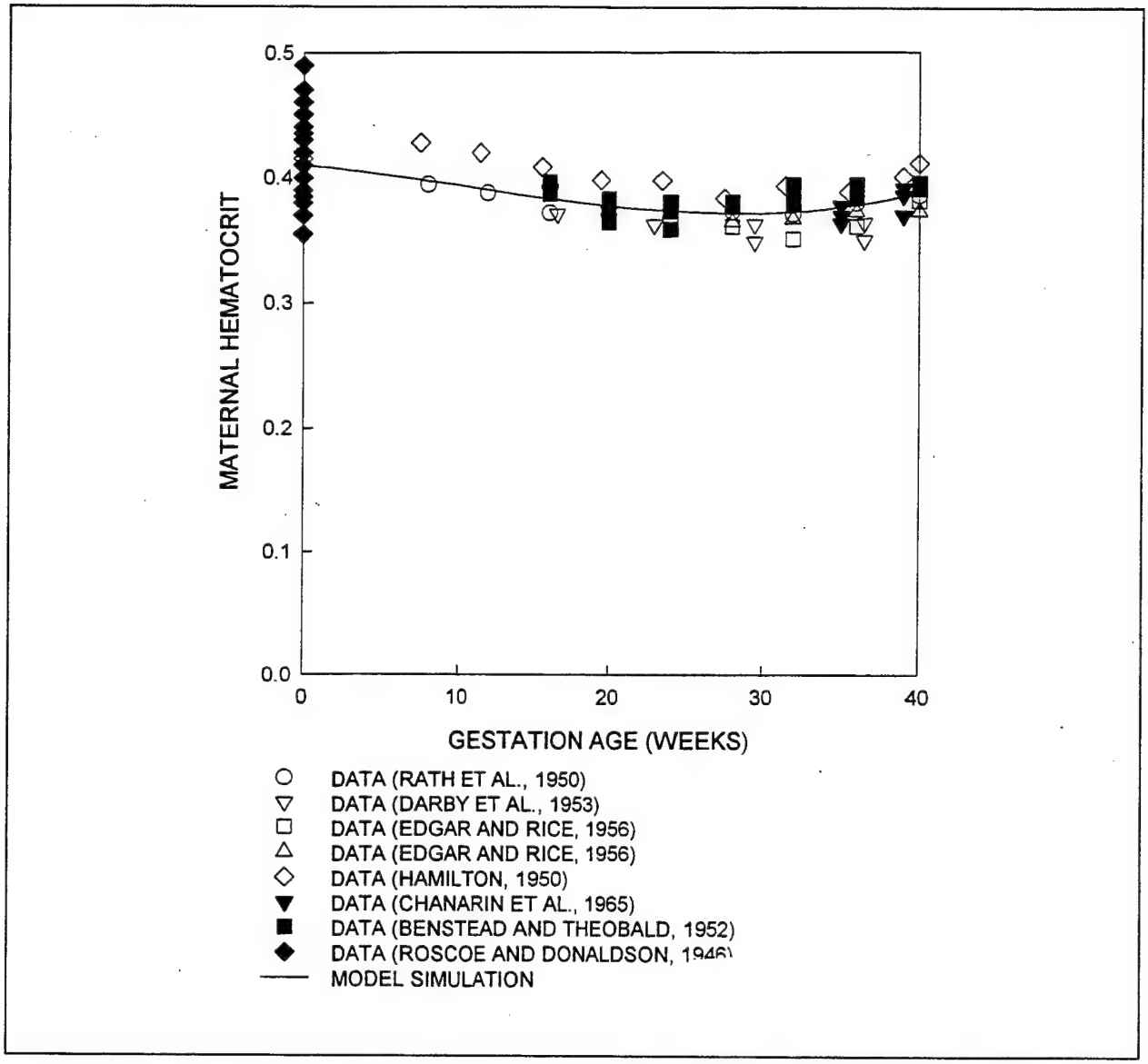


Figure III-8. Change in Hematocrit During Human Pregnancy. The curve shown was judged visually to be the best fit to these data in concert with measurements of plasma and whole blood volumes.

A physiologically-based model of human pregnancy has been developed and partially validated (Andriot and O'Flaherty, 1994). This model includes explicit consideration of changes in bone metabolism (formation and resorption) during pregnancy. It was used to simulate the changes in maternal blood lead concentrations that might be expected during pregnancy in women with high previous lead exposure, with low previous lead exposure, with lead exposure of long or short duration, and with lead exposure beginning or terminating at the beginning of pregnancy. Examples of the

simulations are shown in Figure III-9. High pre-pregnancy exposure, simulated as air lead concentration 0.0035 mg/m^3, drinking water lead concentration 0.004 mg/L, and dietary lead intake by adults 400 μg/day, leads to a maternal blood lead of about 16 μg/dL at the start of pregnancy at age 28. Low lead exposure was simulated as air lead 0.002 mg/m^3, drinking water lead 0.002 mg/L, and dietary lead intake by adults 170 μg/day, leading to a blood lead of about 6 μg/dL at age 28. Simulated maternal blood increases slightly, then decreases and increases again during pregnancy (Figure III-9 – A and B). Reversing exposure conditions at the beginning of pregnancy has a significant impact on predicted maternal blood lead concentrations during pregnancy (Figure III-9 – C and D). Figure III-9D illustrates one exposure scenario that is actually predicted to lead to a net decrease in blood lead concentration during pregnancy. Varying exposure at different ages prior to pregnancy also has an impact on predicted maternal blood lead concentrations during pregnancy (Figure III-9 – E and F).

DISCUSSION

Inclusion in physiologically-based models of age- or time-dependence of anatomic and physiologic parameters requires a fairly full set of experimental observations of these parameters. However, once the data are available their integration into the models is straightforward. Incorporation of a temporal dimension greatly increases the flexibility and applicability of the models without displacing them from their basis in biological reality. Within an appropriate model, exposure estimates can be extrapolated across time as well as across dose, route, and species. It should be noted that temporal behavior is frequently different from one species to another. For example, the average rate of growth from conception to maturity, the temporal pattern of growth to maturity, and the timing of sexual maturity relative to growth are all strongly species-dependent (McCance and Widdowson, 1986). Thus, extrapolation across species is subject to the additional requirement that the key temporal behaviors are known for both species. Such extrapolations are subject to the uncertainties that attend all modeling endeavors. Nonetheless, at their best they offer a means of reconciling and integrating observations made under a wide variety of conditions into a coherent broad picture of kinetic behavior.

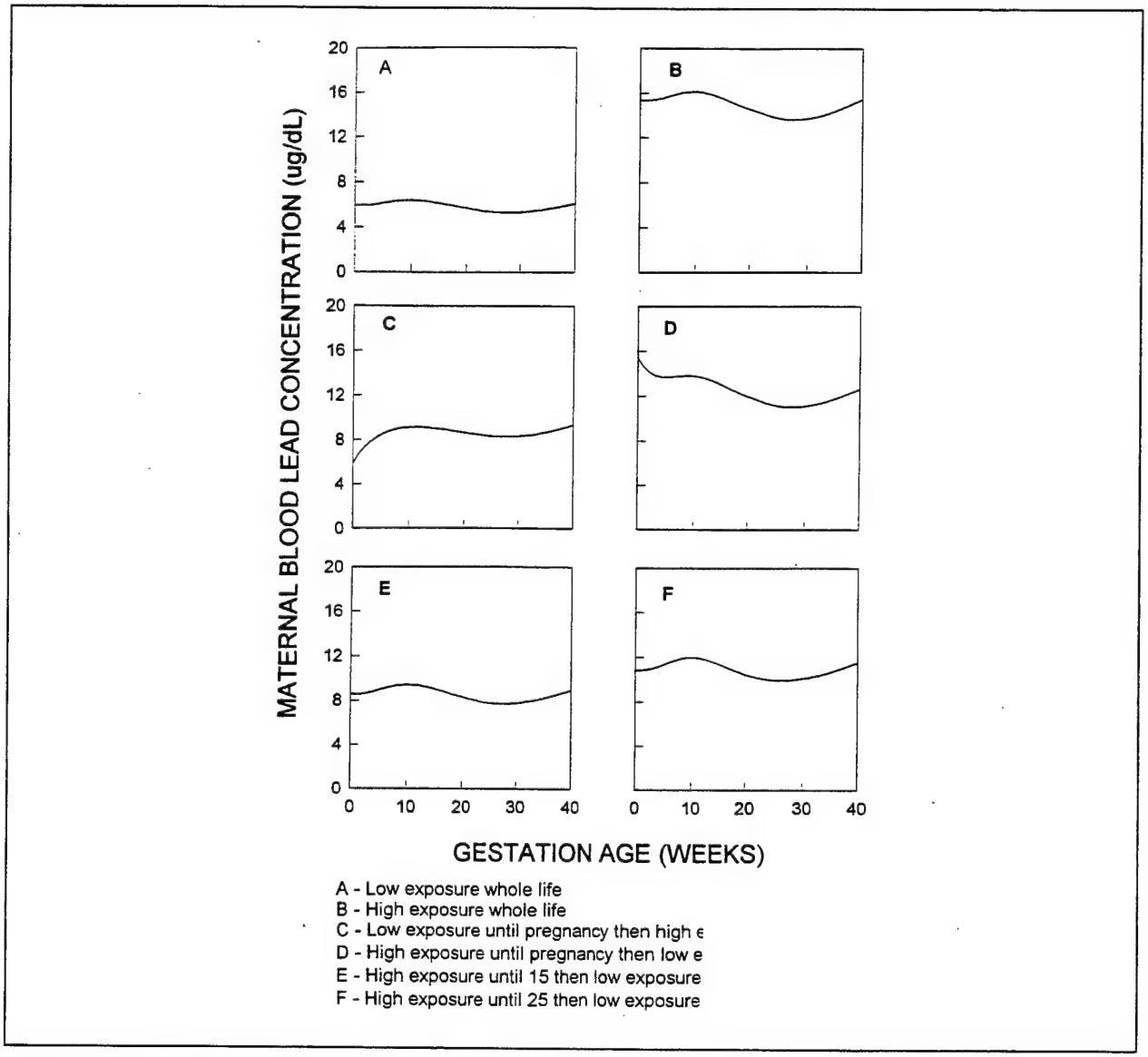


Figure III-9. Predicted Changes in Maternal Blood Lead Concentration during Pregnancy with Varying Exposure Conditions.

REFERENCES

Alexander, F.W. (1974). The uptake of lead by children in differing environments. *Environ. Health Perspect.* **7**:155–159.

Andriot, M.D. and O'Flaherty, E.J. (1994). Changes in bone resorption during pregnancy and incorporation into a physiologically-based pharmacokinetic (PBPK) model for lead. *The Toxicologist* **14**:37.

Freeman, G.B., Johnson, J.D., Killinger, J.M., Liao, S.C., Feder, P.I., Davis, A.O., Ruby, M.V., Chaney, R.L., Lovre, S.C., and Bergstrom, P.D. (1992). Relative bioavailability of lead from mining waste soil in rats. *Fund. Appl. Toxicol.* **19**:388–398.

Hytten, F.E. and Leitch, I. (1971). *The Physiology of Human Pregnancy.* Blackwell Scientific Publications, Oxford.

Johnson, T.R., Moore, W.M., and Jeffries, J.E. (1978). *Children Are Different.* Ross Laboratories Division of Abbott Laboratories, Columbus, OH, p.99.

McCance, R.A. and Widdowson, E.M.. (1986). Glimpses of comparative growth and development. Chapter 7 in Falkner, F. and Tanner, J.M., *Human Growth: A Comprehensive Treatise*, Second Edition, Plenum Press, New York.

O'Flaherty, E.J. (1991). Physiologically-based models for bone-seeking elements II. Kinetics of lead disposition in rats. *Toxicol. Appl. Pharmacol.* **111**:313–331.

Pounds, J.G., Long, G.J., and Rosen, J.F. (1991). Cellular and molecular toxicity of lead in bone. *Environ. Health Perspec.* **91**:17–32.

Rabinowitz, M.B., Kopple, J.D., and Wetherill, G.W. (1980). Effect of food intake and fasting on gastrointestinal lead absorption in humans. *Am J. Clin. Nutr.* **33**:1784–1788.

Watson, W.S., Morrison, J., Bethel, M.I.F., Baldwin, N.M., Lyon, D.T.B., Dobson, H., Moore, M.R., and Hume, R. (1986). Food iron and lead absorption in humans. *Am. J. Clin. Nutr.* **44**:248–256.

Ziegler, E.H., Edwards, B.B., Jensen, R.L., Mahaffey, K.R., and Fomon, S.J. (1978). Absorption and retention of lead by infants. *Pediat. Res.* **12**:29–34.

CONSIDERATION OF TEMPORAL TOXICITY CHALLENGES CURRENT DEFAULT ASSUMPTIONS[1]

Annie M. Jarabek
Environmental Criteria and Assessment Office (MD-52)
U.S. Environmental Protection Agency, Research Triangle Park, NC 27711

ABSTRACT

According to the 1983 NAS paradigm that serves as the basis for current health risk assessment procedures, risk characterization requires the comparison of an exposure estimate against a dose–response estimate. The types of exposure scenarios required under various regulations can be categorized as acute, subchronic, and chronic. Toxicity testing studies can also be so categorized, but such categories are defined by the exposure duration and not the underlying mechanism of action or an appropriate dose metric. Considerations of underlying mechanisms and temporal relationships of toxicity challenge current default assumptions and extrapolation approaches for derivation of dose–response estimates. This paper discusses the duration adjustments used in current health risk assessment procedures and highlights the attendant assumptions. Comprehensive dosimetry model structures integrate mechanistic and temporal determinants of the exposure-dose–response continuum. Analysis of dosimetry model structures is proposed as a way to identify key parameters for development of alternative default duration adjustment procedures.

The various environmental and occupational regulatory statutes and implementation activities under such laws as the Clean Air Act Amendments of 1990 (CAAA), the Safe Drinking Water Act (SDWA), the Clean Water Act (CWA), the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), the Resource Conservation and Recovery Act (RCRA), and the Occupational Safety and Health Act (OSHA) require risk characterization and risk management of exposure scenarios that range in duration from a few minutes to lifetime. The 1983 NAS/NRC report on risk assessment and risk management presented a paradigm[a] for this process that serves as the basis of most health assessment procedures and regulatory programs in various federal agencies (NRC, 1983). In order to characterize health risk for these different scenarios so that risk management decisions may be made, dose–response estimates for toxicity that are comparable to these exposure scenarios must be

[1] The views expressed in this paper are those of the author and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency. The U. S. Government has the right to retain a nonexclusive royalty-free license in and to any copyright covering this article.

derived. The definition of comparability between exposure and toxicity estimates, however, has usually been based on the comparability of the exposure duration of the objective exposure scenario to that of the experimental exposure in the laboratory test species (or to various dose surrogates commonly used in occupational epidemiology). In most cases, these definitions do not take into account the mechanistic and temporal determinants of the toxicity nor account for the species differences in such determinants.

This paper outlines the current dose–response procedures typically used for noncancer toxicity of various durations. The assumptions underlying current procedures for duration extrapolation are discussed and evaluated with consideration of potential mechanistic and temporal determinants of toxicity. Since toxicity depends on the magnitude, duration, and frequency of exposure; which in turn can be affected by the timing of exposure; determination of the appropriate dose metric and duration extrapolation should thus be dependent on the mechanism of toxicity. Dosimetry models incorporate mechanistic determinants of chemical disposition in order to characterize the relationship between exposure concentration and target tissue dose. Because these disposition determinants include both concentration and time dependent processes, analysis of dosimetry model structures is proposed as a way to identify key parameters and to define limiting conditions for development of alternative default duration extrapolation procedures.

EXPOSURE SCENARIOS AND TYPES OF TOXICITY DATA AVAILABLE

As mentioned above, various regulatory statutes and implementation activities require health risk characterization of different exposure scenarios as the basis of risk management programs. As shown in Table III-2, these exposure scenarios range from minutes to lifetime, and are often categorized as acute, subchronic, and chronic. Certain of these scenarios have default assumptions incorporated in their definition. For example, default consumption values of 24-h continuous inhalation exposure (at a rate of 20 m^3/day) and 2 L/day for water intake are assumed for a 70 kg person (male). Eight-hour, time-weighted averages are often used as exposure surrogates for occupational scenarios. A daily (24-h) average exposure may be used as an exposure surrogate for "acute" ambient exposures whereas the annual average is calculated as a surrogate for "chronic". "Lifetime" or "chronic" exposures for humans are assumed to be 70 yr, and 10% of this lifetime (7 yr) defines the lower cut-off for "subchronic" exposures. Exposures are usually assumed to be at a constant concentration whereas the actual exposure is a profile dependent on numerous factors such as production volume, stack height, meteorology, and human activity patterns.

Table III-2. Exposure Scenarios Requiring Risk Characterization.

Acute	15-min Occupational TWA* Ceiling Level
	1-h Emergency Response Planning Guidelines
Subchronic	Intermittent Start-Up/Shut-Down Processes
	Peridoic Contaminations
Chronic	8-Hour Occupational TWA Exposure Limits for "Working Lifetime"
	Ambient Exposures for "Lifetime"

*TWA = To,e-weighted average.

Experimental exposures to animals are typically divided into four categories: acute, subacute, subchronic and chronic, as shown in Table III-3. Acute exposure is defined as an exposure to a chemical for less than or equal to 24 h. Although usually for a single administration (e.g., 4 h), repeated exposures are sometimes given within the 24-h period. Repeated or continuous exposures are also divided into subacute, subchronic, and chronic categories. Subacute refers to repeated or continuous exposure to a chemical for 1 mo or less (e.g., a 14-day range finding study). Subchronic refers to repeated or continuous exposure for 1 to 3 mo, usually a 90-day study. Chronic refers to repeated or continuous exposure for longer than 3 mo, most commonly a 2-yr bioassay in rodents.

Table III-3. Exposure Scenarios Requiring Risk Characterization.

Acute	1–24 h single Inhalaiton Exposures
	Single (or few) Oral Administrations
Subacute	14-Day Range-Finding Exposures
Subchronic	90-Day Exposure Studies
Chronic	2-Year Bioassays

These categories are defined based on the duration of the exposure and in the absence of any consideration of mechanisms of toxicity or its temporal aspects. Generally, acute toxicity data are used as the basis for derivation of acute toxicity dose-response estimates that are used to compare against acute exposure estimates for risk characterization. Likewise, chronic bioassay data (or subchronic data with application of an uncertainty factor for the affect of duration) are used as the basis for derivation of chronic toxicity dose–response estimates for characterization of lifetime ambient exposure scenarios.

Thus, a fundamental assumption of these approaches is that toxicity across different species is a function of lifetime fraction (chronologic) exposed (e.g., a 2-yr "lifetime" bioassay in the rodent is equivalent to a 70-yr human "lifetime" for the purposes of chronic health risk characterization).

The attendant uncertainties and default assumptions of the dose–response estimate should be evaluated in context with those of the exposure estimate (e.g., assumptions of fate and transport modeling or type of sampling and averaging time of a measured exposure) to ascertain whether the two are appropriate to integrate. Table III-4 provides a comparison of different assumptions and derivation methods inherent in some common risk assessment and risk management estimates. The intended use of a dose–response or risk management estimate influences its derivation (Jarabek and Segal, 1993). The assumptions and uncertainties of the risk characterization components (dose–response and exposure assessments) must be explicitly communicated to the risk management arena for application to intended scenarios. Often dose–response estimates are compared inappropriately with risk management or regulatory values that are intended for different exposure scenarios and populations or that are derived using additional considerations such as control technology. Because of these differences, the remainder of this paper will discuss only procedures for dose–response estimation.

Table III-4. Comparison of Exposure Limits.

Organization/ Exposure Limit	NAS Paradigm	Objective Exposure Scenario	Effect Severity	SF or UF[1]	Population	Derivation/ Data Base
"Less than Lifetime" Exposure Limits						
ACCIH TLV-STEL[2]	Management	15 min Time-Weighted Average Exposure That Should Not Be Repeated More than 4 Times per Day	Protect Against Irritation, Chronic or Irreversible Tissue Damage, or Narcosis of Sufficient Degree to Increase Chances of Accidental Injury, Impair Self-Rescue or Reduce Work Efficiency	Minimal SF Used; No Systematic Application	Health Worker	Based on best available information from industrial experience, experimental human and animals studies (human data preferred). No systematic basis; derived by expert committee.
AIHA ERPG[3]-3	Management	1-h Exposure	Protect Against Life-Threatening Effects	SF	General Population Living in Immediate Area of Release	Acute toxicity data preferred. Based upon most sensitive endpoint from human or animal data. All endpoints considered. Methods vary on a case-to-case basis.
AIHA ERPG[3]-2	Management	1-h Exposure	Protect Against Irreversible of Other Serious Health Effects That Could Impair Ability to Take Protective Action	SF	General Population Living in Immediate Area of Release	Same as ERPG-3
AIHA ERPG[3]-1	Management	1-h Exposure	Protect Against Mild, Transient Adverse Health Effects	SF	General Population Living in Immediate Area of Release	Same as ERPG-3
COT EEGL[4]	Management	1- and 24-h Exposures	Revesible Effects Acceptable, (e.g., headache, irritation, CNS effects)	Generally No (unless confidence in data base is low or chemical is a carcinogen)	Military Personnel, Assumed to Be Healthy and Relatively Homogeneous	Based on most sensitive endpoint (NOAEL or LOAEL) from human or animal toxicity data (acute toxicity data preferred). All endpoints considered.

Table III-4. Comparison of Exposure Limits.

Organization/ Exposure Limit	NAS Paradigm	Objective Exposure Scenario	Effect Severity	SF or UF[1]	Population	Derivation/ Data Base
COT SPEGL[5]	Management	1- and 24-h Exposures	Revesible Effects Acceptable, (e.g., headache, irritation, CNS effects)	SF of 2–10 Applied to EEGL to Protect More Sensitive Sub-populations (SF = 2) or Fetuses or Newborns (SF = 10)	General Population	EEGL divided by a factor of 2–10 to protect more sensitive subpopulations.
COT CEFL[6]	Management	90-Day Exposures	Revesible Effects Acceptable, (e.g., headache, irritation, CNS effects)	SF of 10–100 Applied to EEGL Based on Pharmaco-kinetics (i.e., ability to be rapidly biotranformed or to bioaccum-ulate)	General Population	EEGL divided by a factor of 10–100 to account for pharmacokinetic considerations.
"Lifetime" Exposure Limits						
AGGIH TLV-TWA[7]	Management	8 h/Day; 40 h/Week for a Working Lifetime (40 years)	No Adverse Effect	Minimal SF Used; No Systematic Application	Nearly all Workers; Personal Protective Equipment May Be Factored	Based on best available information from industrial experience, experimental human and animal studies (human data preferred). No systematic basis; derived by expert committee.

Table III-4. Comparison of Exposure Limits.

Organization/ Exposure Limit	NAS Paradigm	Objective Exposure Scenario	Effect Severity	SF or UF[1]	Population	Derivation/ Data Base
NIOSH REL[8]	Management	Up to 10 h/Day; 40 h/Week; Undefined Working Lifetime Duration; Appropriate Control and Surveillance Methods	No Adverse Effect	Minimal SF Used; No Systematic Application	Nearly all Workers; Personal Protective Equipment May Be Factored	Based on best available information from industrial experience, experimental human and animal studies (human data preferred). No systematic basis; derived by expert committee.
OSHA PEL[9]	Management	8 h/Day 40 h/Week; 45 Year Working Lifetime Duration; Appropriate Control and Surveillance Methods	Protect Worker Against a Wide Variety of Health Effects That Could Cause Material Impairment of Health or Functional Capacity	Same As Above	Same As Above	Same as above. In addition, technological feasibility is considered in establishing a PEL.
ASTSDR MRL[10]	Dose-Response	24 h/Day 70 Year	NOAEL	UF	General Population Including Susceptible	Occupational, experimental human and animal.

Adapted from: Jarabek and Segal (1993) and Jarabek (1994).

[1]SF = Safety Factor; UF = Uncertainty Factor consistently used for explicit extrapolations applied to data.

[2]ACGIH = American Conference of Governmental Industrial Hygienists; TLV-STEL = Threshold Limit Value - Short-Term Exposure Level.

[3]AIHA = American Industrial Hygiene Association; ERPG = Emergency response planning guideline

[4]COT = Committee on Toxicology of NAS; EEGL = Emergency exposure guidance level.

[5]COT = Committee on Toxicology of NAS; SPEGL = Short-term population exposure guidance level

[6]COT = Committee on Toxicology of NAS; CEGL = Community Exposure Guidance Level

[7]TLV-TWA = Threshold-Limit Value ¾ Time-Weighted Average

[8]NIOSH = National Institute for Occupational Safety and Health; REL = Recommended Exposure Level.

[9]OSHA = Occupation Safetly and Health Agency; PEL = Peak Exposure Level.

[10]ATSDR = Agency for Toxic Substances and Disease Registry; MRL = Minimum Risk Level.

DEFAULT DURATION EXTRAPOLATION FOR NONCANCER TOXICITY

Current procedures for dose–response estimation attempt to match the durations of the exposures that are the basis of the toxicity data with the anticipated human exposure scenario. For oral exposures, this assumption applies whether or not the dose was administered as ppm in water, in diet, or via gavage. When the exposure duration of the laboratory animal toxicological study does not match that of the objective human exposure scenario, a linear prorated adjustment of the exposure concentration is typically performed. The default duration adjustments are shown below for acute and chronic inhalation exposures.

Acute (≤24 h) Exposures

$$EL_{ADJ}\ (ppm) = EL\ (ppm) \times D\ (h)/H\ (h) \qquad (1)$$

where EL_{ADJ} is the effect level (ppm), such as a No-Observed-Adverse-Effect-Level (NOAEL) or Lowest-Observed-Adverse-Effect-Level (LOAEL), adjusted for duration of experimental regimen; EL is the experimental exposure level (ppm); D is the experimental exposure duration (h), and H is the objective human exposure duration (h). Twenty-four hours is used by the Agency for Toxic Substances and Disease Registry (ATSDR) as the default for the objective acute human exposure duration (H).

Chronic (2-yr bioassay) Exposures

$$EL_{ADJ}\ (ppm) = EL\ (ppm) \times D\ (h/24h) \times W\ (days/7days) \qquad (2)$$

where W is the weekly frequency of exposure in days/7 days.

Thus, the default duration adjustments assume that exposure concentration is equivalent to inhaled dose. Further, it is also assumed that toxicity is linearly related to the product of C × t so that equivalent products cause the same toxicity. A notable exception to these duration adjustments is that of developmental toxicity. No duration adjustment is applied to effect levels for this noncancer toxicity. The rationale is that because developmental toxicity can occur within any time window of the gestational period, duration extrapolation is inappropriate.

The assumption in Equations 1 and 2, that the resultant human exposure concentration should be the concentration times time (C × t) equivalent of the experimental animal exposure level, is based on "Haber's Law"[b]. According to this "law", a constant, in this case a fixed effect (i.e., a constant severity and/or incidence) level is related to exposure concentration and duration as:

$$EL = aC^n \times t\,, \qquad (3)$$

where EL is the fixed effect level, a is a coefficient defined empirically, C is the exposure concentration, n is an exponent defined empirically, and t is the duration of exposure. Figure III-10 is a schematic illustration of the relationship between exposure concentration and duration to a fixed effect level (EL) assuming "Haber's Law". The relationship is described by a hyperbola whose arms converge asymptotically toward the axes of the coordinates (Bliss, 1940). Because Haber examined only extremely short durations, a C × t relationship appeared to hold because concentration was the dominant determinant of toxicity in that limited time window. Bliss and James (1966) showed that such curves could be extrapolated with minimal error only when the time points in the experiment are located on the asymptotic segments of the curve (i.e., high concentration, acute exposures or low concentration, chronic exposures). The rationale when applied to chronic exposures is that the concentration is low and steady state has been reached and thus duration is the dominant determinant.

"Haber's Law" is related to the logtime-logdosage curve. When the relationship in Equation 3 is plotted on loglog paper, all solutions lie on a straight line. When the exponent, n, is equal to 1, the line passes through the two points (C = 1, t = EL) and (C = EL, t = 1) and has a 45-degree slope. Empirical data have also shown greater or lesser slopes. The smaller the exponent, the steeper the slope. Based on 1-h lethality studies, ten Berge et al. (1986) found that 19/20 substances showed a value for n in Equation 3 to be in the range of 1.0 to 3.5. The one exception had a value of 0.8. If 4-h exposure level is extrapolated to shorter durations using the C × t assumption and a value of 1 for n, the resultant estimate is considerably higher than that estimated using a value of 3.5 for n. Based on this analysis, ten Berge et al. (1986) concluded that estimates of the C × t relationship for derivation of extrapolation procedures should be developed using chemical-specific information. Data to construct logtime-logdosage plots are available for relatively few chemicals and most are lethality data. Establishing an extrapolation procedure based on lethality data may not be appropriate for milder effects such as a NOAEL or LOAEL used in risk assessment, especially for extrapolation to a shorter duration, because mechanistic determinants may be different for severe versus milder effects.

Figure III-11 illustrates an example of the potential inaccuracies in the prorated linearized extrapolation approach to either shorter or longer durations. The vertical dashed lines at 1-h and 8-h indicate typical exposure durations required for estimation. If the single 4-h experimental concentration was extrapolated to a 1-h exposure estimate, assumption of Haber's Law (B) results in an overestimate of exposure when compared to an estimate assuming concentration alone (A) is the dominant determinant of toxicity. Extrapolation to the longer 8-h duration results in the converse relationship

between exposure estimates—in this case the estimate derived assuming Haber's Law (B) is conservative in comparison to the estimate based on concentration as a constant (A).

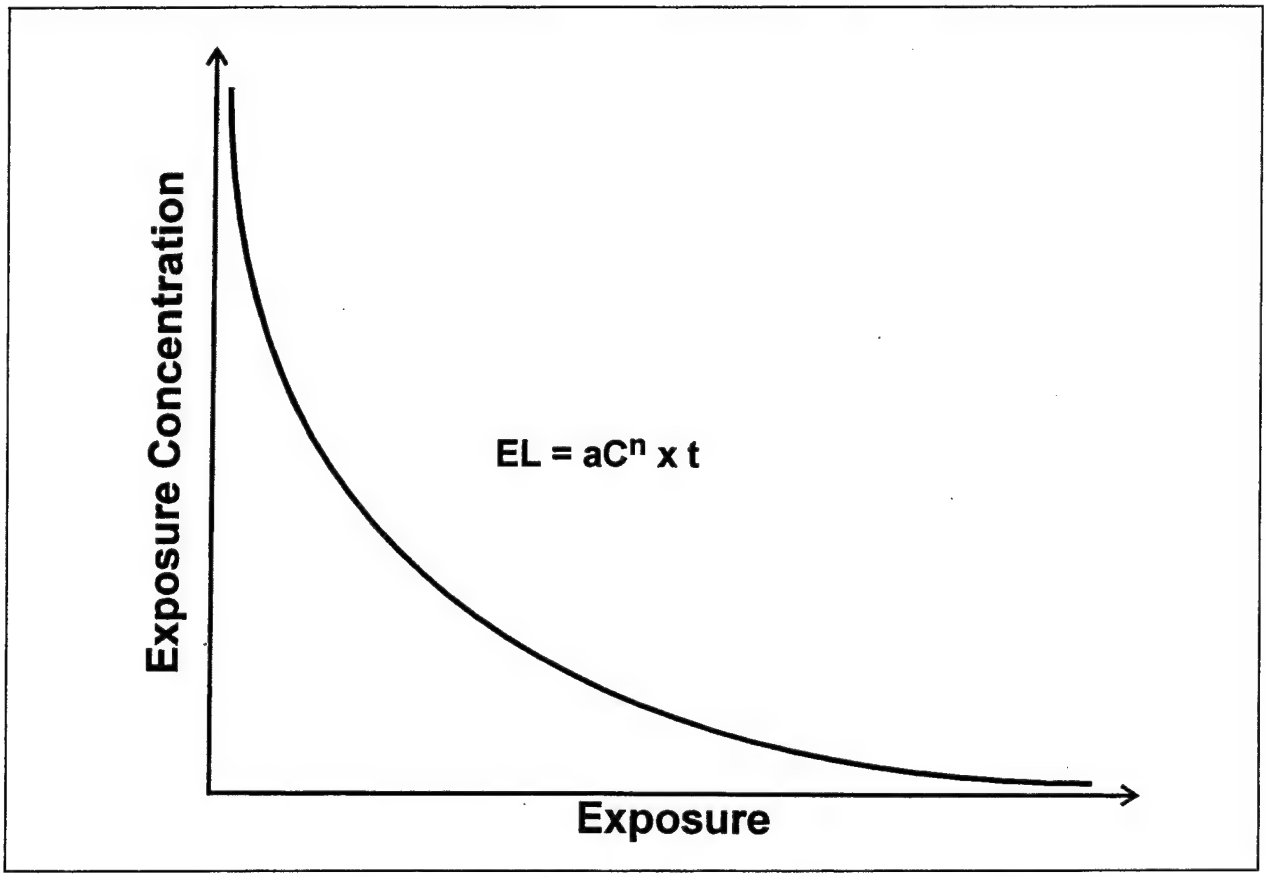


Figure III-10. Schematic of Relationship between Exposure Concentration (C) and Duration (t) to Fixed Effect Level (EL) of Toxicity Assuming "Haber's Law".

Extrapolation of an exposure based on Haber's Law versus keeping concentration a constant regardless of duration could be based on consideration of what the mechanism of action is believed to be. Extrapolation keeping EL = C could be appropriate for irritants, so that no matter the duration, the effective concentration level remains the same. In this case, concentration alone is used as a dose metric. If either the chemical or its damage accumulates with duration, the exposure level given by extrapolation using Haber's Law would be more appropriate because duration (t) is an explicit determinant. Andersen et al. (1987a) suggested that toxicity for most industrially important gases and volatile liquids would probably be related to the area under the blood curve (AUBC) rather than to peak blood concentrations, so that the use of C × t in the absence of sufficient mechanistic data might

be an acceptable way of extrapolation because the AUBC would be a similar estimate. Some dose-response methods qualitatively take mechanistic data into account and caveat the use of the above default duration adjustments (U.S. EPA, 1994). Consideration of cases where the C × t assumption may not hold is encouraged (e.g., when concentration may be the dominant determinant). The recent inhalation RfC for 2-chloro-1,1,1,2-tetrafluorethane (HCFC-124) did not use the duration adjustment because the data suggested the reversible narcotic effect was due to parent compound concentration only and it had a short half-life (U.S. Environmental Protection Agency, 1993). For most effects, however, the "true" dose metric is not determined and in all likelihood, extrapolation for many toxicants should lie somewhere between the two lines.

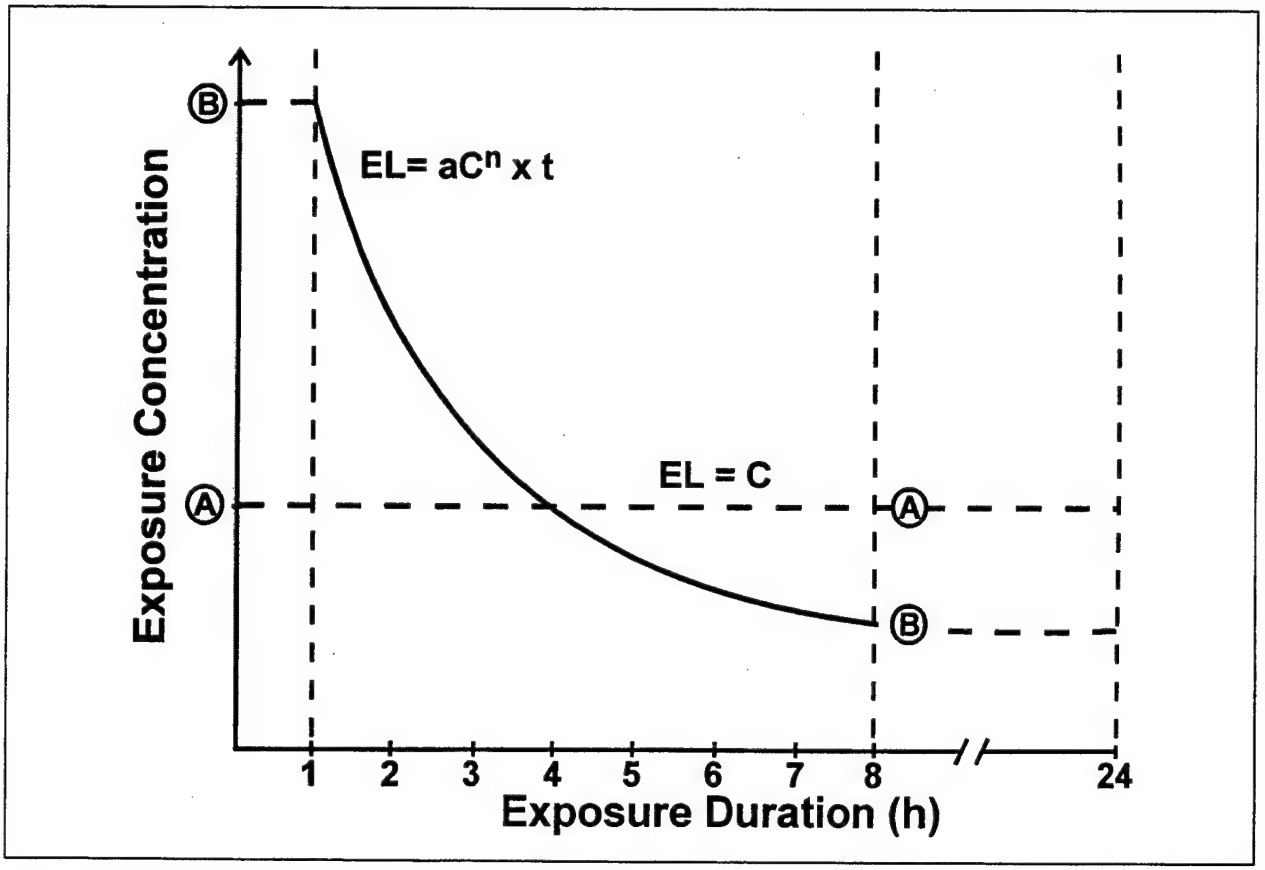


Figure III-11. Schematic of Relationship between Exposure Concentration (C) and Duration (t) to Fixed Effect Level (EL) of Toxicity Assuming "Haber's Law" (solid line) versus Concentration as the Major Mechanistic Determinant (dashed line). Resultant effect levels calculated by extrapolation from a 4-h exposure to a 1-h and 8-h exposure are shown for both.

MECHANISTIC DETERMINANTS AND TEMPORAL ASPECTS OF TOXICITY

Toxicity can depend on the magnitude, duration, and frequency of exposure. Timing in turn can affect these parameters (e.g., different windows within gestation have different susceptibility). Mechanistic determinants of chemical disposition (deposition, absorption, distribution, metabolism and elimination) of a chemical include both time and concentration dependent processes. In general, fractionation of the dose reduces the effect. If detoxifying biotransformation or elimination occurs between successive doses, or if the damage produced is repaired between successive doses, then a single dose may produce more toxicity than that same amount fractionated into many smaller doses given at intervals. Chronic effects occur if the chemical accumulates, if it produces irreversible effects, or if there is insufficient time for the target tissue to recover from the damage within the exposure frequency interval. Acute toxicity may or may not resemble that manifest after prolonged repeated exposures. For many chemicals, the critical toxic effects following a single high concentration exposure are quite different from those produced by repeated low-level exposure (e.g., the acute toxic manifestation of high concentration benzene exposure is central nervous system depression but chronic low-level exposure can result in blood dyscrasias and leukemia). Acute exposure can also produce delayed toxicity. Conversely, chronic exposure to a toxic agent may produce some immediate (acute) effects after each exposure in addition to the long-term chronic effects. Thus, to truly characterize the toxicity of a specific chemical, information is needed not only on acute and chronic effects but also for exposures of intermediate duration.

The choice of an appropriate measure of "dose" must be defined by the nature of the pathogenesis process (i.e., defined according to the mechanism of action) for the effect under consideration. For example, the appropriate dose metric for the CNS depression of acute high concentration benzene exposure could be the parent compound blood concentration, whereas the area under the tissue concentration curve for toxic metabolites would be more appropriate to characterize the erythroid precursor pertubations of chronic low-level exposures. Examples of other potential dose metrics are provided in Table III-5 and illustrated in Figure III-12. Because tissue dose of the putative toxic moiety for a given response is not always proportional to the applied dose of a compound, emphasis has been placed on the need to clearly distinguish between exposure concentration and dose to critical target tissues. The term exposure-dose–response assessment has been recommended as more accurate and comprehensive (Andersen et al., 1992).

The process of determining the exposure-dose–response continuum is achieved by linking descriptions of the mechanisms of critical biological factors that regulate the occurrence of a particular process and the nature of the interrelationships among these factors. The iterative process of linking descriptions at various stages along the continuum is shown in Figure III-1? It is ultimately desirable to have a comprehensive biologically-based dose response model that incorporates the mechanistic determinants of chemical disposition, toxicant-target interactions, and tissue responses integrated into an overall model of pathogenesis. Dosimetry models can be linked to pharmacodynamic models that address the mechanistic determinants of the toxicant-target tissue interaction and tissue response, respectively. Biologically-based dose response models refine the designation of response. The tissue dose is linked to determinants of target-tissue interaction (e.g., critical mechanistic events such as cytotoxicity and rebound cellular proliferation), which, in turn, may then be related via other mechanisms to the ultimate production of lesions or functional changes that are typically defined as the disease (pathogenesis) outcome. To the extent that these intermediate events are explanatory of the disease outcome, they can be used to quantitate important nonproportionalities or as replacement indices of the response function. For example, the amount of cytotoxicity from a low-level exposure to a chemical known to cause cellular proliferation and subsequent neoplasia could be used to evaluate risk using mechanistic models rather than estimation of risk based on tumors from high concentration exposures.

Table III-5. Potential Dose Metrics.

Exposure Concentration of Parent Chemical
Blood Concentration of Parent Chemical
AUBC of Parent chemical
Tissue Concentration of Parent Chemical
AUTC of Parent Chemical
Tissue Concentration of Metabolite AUTC of Stable Metabolite
AUTC of Stable Metabolite
AUTC of Reactive Metabolite

AUBC = Area under blood concentration curve.
AUTC = Area under tissue concentration curve.

Current dose–response assessment methods are essentially based on characterization of the exposure-dose–response continuum at the first, "black box" level and necessarily incorporate large uncertainty factors to ensure that the estimates are protective in the presence of data gaps that are often

substantial. Use of "Haber's Law" can be viewed as falling in this first "black box" tier. Interestingly, Hayes (1975) restated the relationship of "Haber's Law" in recognition of its limited applicability to address dosimetry considerations as:

$$[(CVm) - De]tR/w = D \quad (4)$$

where D is the dosage (mg/kg) received during time, t; C is the concentration of toxicant (mg/m^3), Vm is the minute volume rate of respiration (m^3/min), De is the detoxification rate (mg/min), t is the time (min) of exposure, w is body weight (kg) and R is the retention coefficient expressed as a decimal fraction. The equation shows that a sufficiently high rate of detoxification would negate prolonged exposure to a sufficiently low concentration. It thus expresses quantitatively the limitation on the rule when applied to easily detoxified materials. It is also seen in Equation 4 that the dosage, D, is not necessarily a constant for all combinations of concentration and time that produce the same effect, since the detoxification rate and perhaps the retention coefficient may vary with dosage.

Equation 4 is actually an attempt to account for potential mechanisms of toxicity and a dose other than the exposure concentration as a metric. Unfortunately, most of the parameters in Equation 4 are not determined routinely in toxicological studies nor would they be available for humans. Since the formulation of this equation, dosimetry models[c] have evolved into particularly useful tools for predicting chemical disposition differences between species.

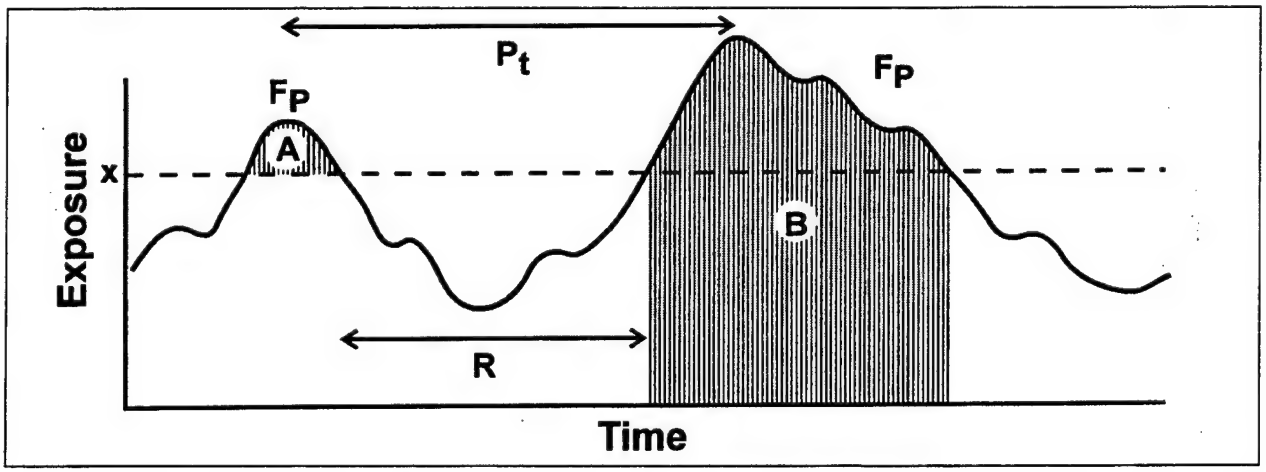


Figure III-12. Effective Exposure Causing Toxicity Depends on the Magnitude, Duration, and Frequency of Exposure. Potential Dose Profile Metrics are Illustrated. A = area over the threshold, ×. B = summation of all dose when × is exceeded. P_t = time between peaks over ×. R = respites between peaks over x. F_p, the frequency of peaks over ×. Integrated dose and average dose could also be calculated.

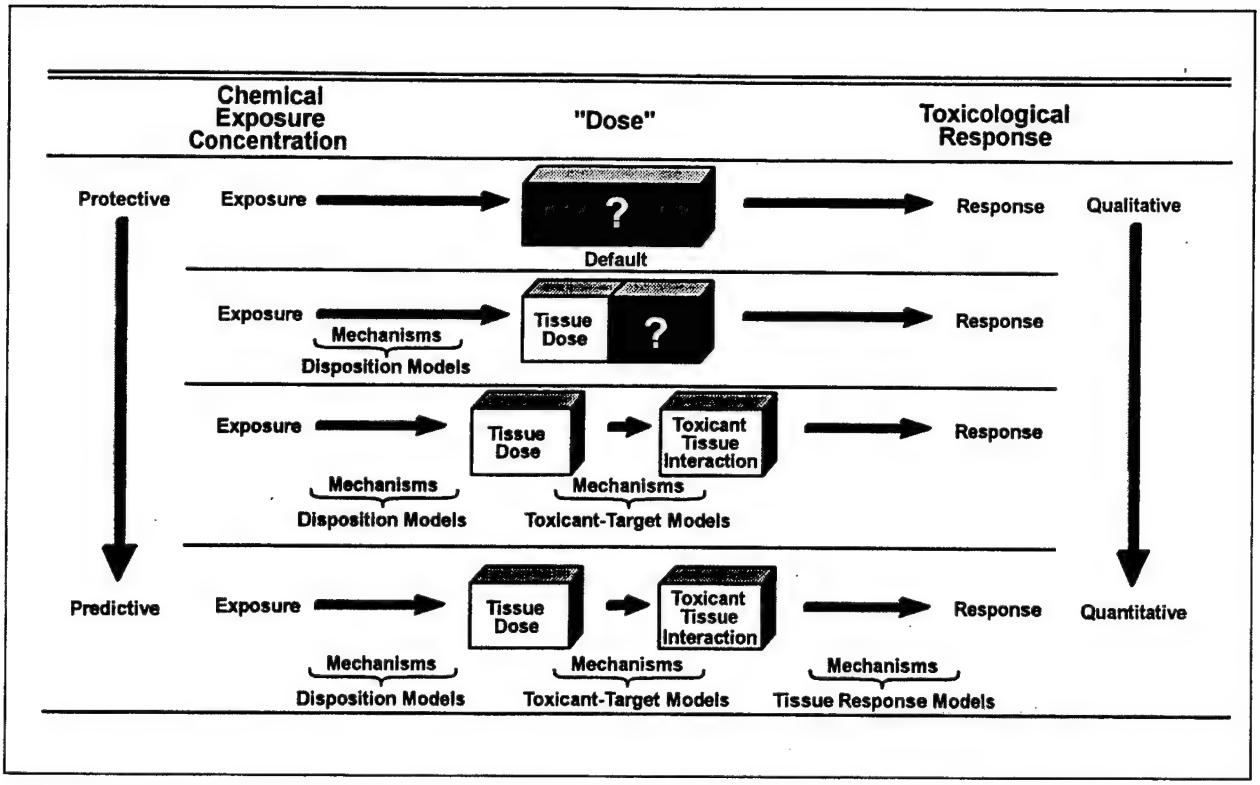


Figure III-13. Schematic Characterization of Comprehensive Exposure-Dose–Response Continuum and the Evolution of Protective to Predictive Dose–Response Estimates. Adapted from Andersen et al. (1992).

APPLICATION OF DOSIMETRY MODELS

Dosimetry models that account for mechanistic determinants of the disposition of a parent compound and/or its metabolites, such as PBPK models, have been useful in describing the relationships between exposure concentration and target tissue dose. Because pharmacodynamic data (data on toxicant-target tissue interactions including differences in response due to sensitivity) are the least available, the majority of dosimetry models have restricted structures to describe chemical disposition. Scaling of mechanistic parameters, such as metabolic rates, provides for accurate extrapolation to humans.

Default dosimetry adjustments using a limited number of key parameters and based on mathematical reduction of more comprehensive dosimetry model structures have been developed for different types of inhaled chemicals (particles and various categories of gases) (U.S. EPA, 1994). Use of these default dosimetry adjustments for interspecies extrapolation has moved the U.S. EPA's

inhalation reference concentration (RfC) methods to the second tier within the framework shown in Figure III-13 (Jarabek, 1995). Because the mechanistic determinants of chemical disposition (deposition, absorption, distribution, metabolism, and elimination) include both time and concentration dependent processes, similar analysis of dosimetry model structures to identify key parameters and processes may serve to provide alternatives to the duration extrapolation based on the C × t assumption of "Haber's Law" (Jarabek and McDougal, 1993). Figure III-14 outlines various parameters and processes that determine the dominant mechanisms at each interface for progression from exposure to response. The parameters outlined have been incorporated in various mathematical models for specific chemicals.

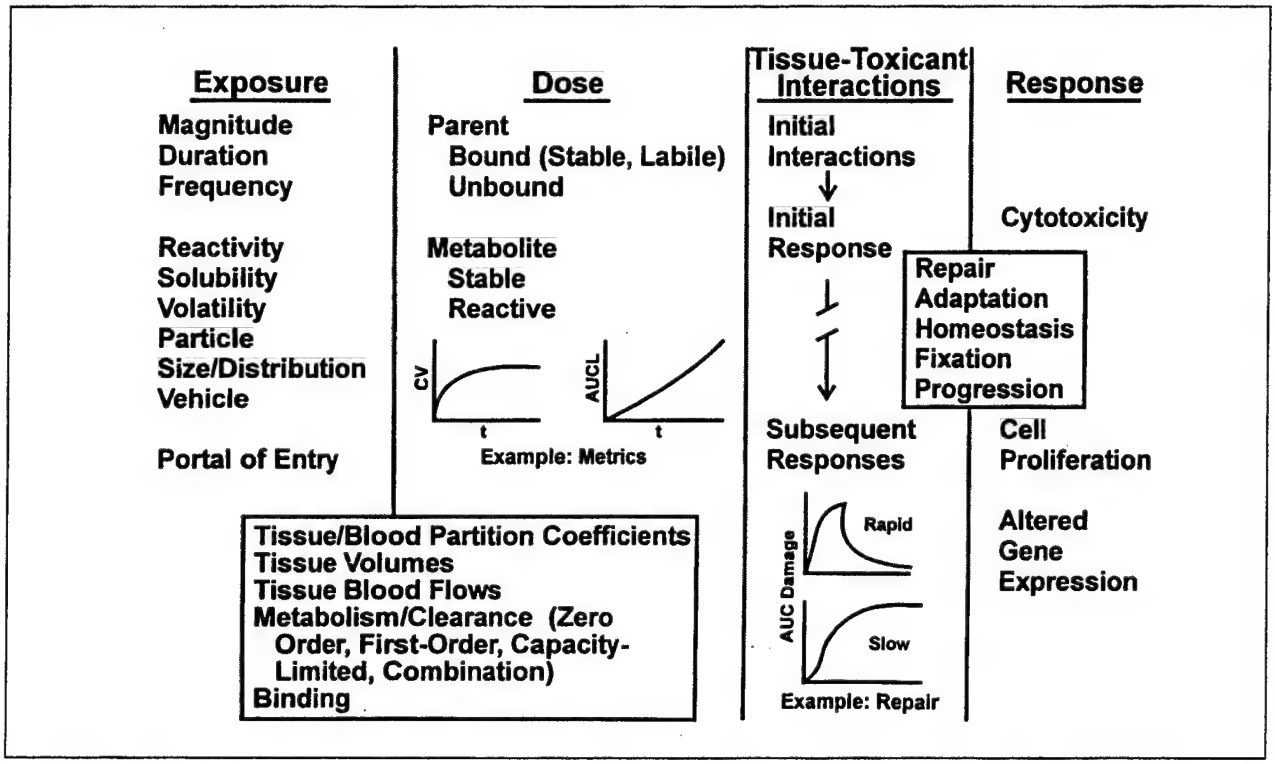


Figure III-14. Schematic of Parameters Important to Defining Interfaces of Exposure-Dose-Continuum.

Figures III-15 and 16 illustrate model simulations of the rate of metabolite formed per gram liver tissue via the mixed-function oxygenase system at each of 3 different C × t products for dichloromethane (DCM) and perchloroethylene (PERC) (Jarabek and McDougal, 1993). Each of the lines connects output from 7 different simulations which have an equivalent C × t exposure product. For example, a 0.5-h exposure at 400 ppm, a 4-h exposure at 50 ppm, and an 8-h exposure at 25 ppm

are simulations which have an equivalent C × t exposure product of 200 ppm-hrs. If "Haber's Law" held, the plot of equivalent C × t products versus t would be a straight horizontal line.

DCM and PERC were chosen because they differ in both key physicochemical parameters (e.g., fat:blood partition coefficients of 19.4 versus 121.0 for DCM and PERC, respectively) and metabolic parameters (e.g., Vmax of 11.54 versus 0.180 mg/h/kg, for DCM versus PERC, respectively). For DCM, concentration is the dominant factor on the rate of metabolism since this chemical has the greater Vmax. At 50 ppm the system is not yet saturated. For PERC, time is the more dominant factor on this dose metric since essentially all three C × t products are above saturation. Differences in other dose metrics, such as venous concentration or area under the liver curve are also exhibited (Jarabek and McDougal, 1993). Similar differences in profiles could be anticipated for oral exposures as dosimetry models have shown major differences in resultant dose metrics after gavage in oil versus water and in comparison to administration in drinking water (Corley and Reitz, 1990).

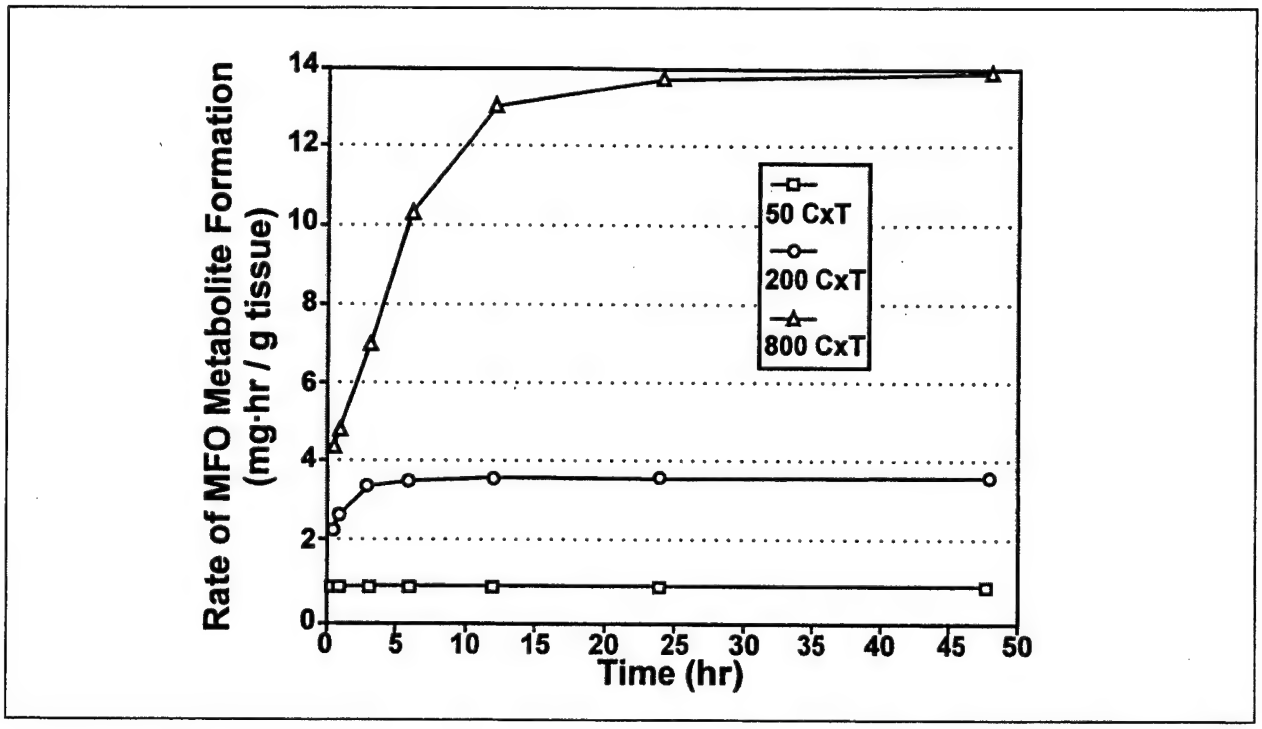


Figure III-15. Model Simulations of the Rate of Metabolite Formed per Gram Liver Tissue via the Mixed Function Oxygenase System in the Rat Simulated at Each of Three Different C × t Exposure Products for DCM. Each of the lines connects output from 7 different simulations which have equivalent C × t exposure products (e.g., a 0.5-h exposure at 400 ppm, a 4-h exposure at 50 ppm, and an 8-h exposure at 25 ppm). The PBPK model used was that published by Andersen et al. (1991). Parameter values used are available elsewhere (Jarabek and McDougal, 1993).

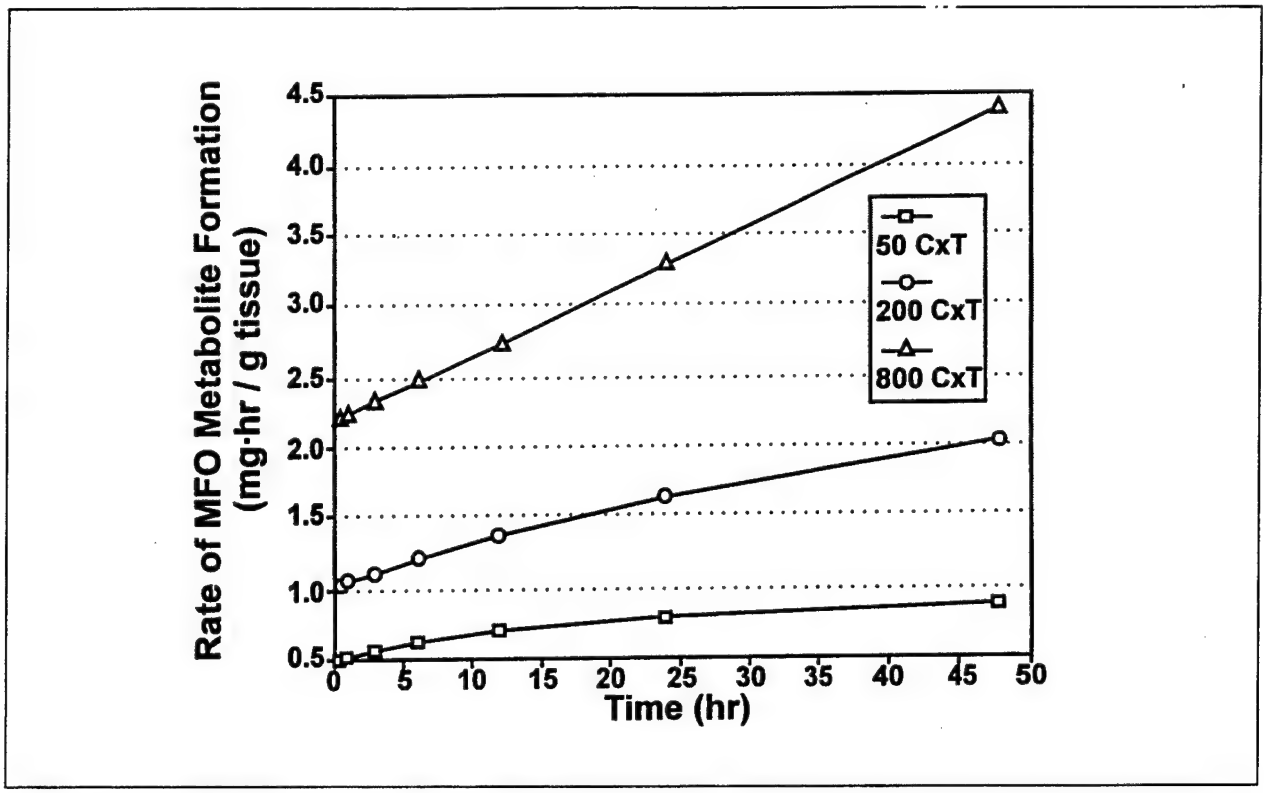


Figure III-16. Model Simulations of the Rate of Metabolite Formed per Gram Liver Tissue via the Mixed Function Oxygenase System in the Rat Simulated at Each of Three Different C × t Exposure Products for PERC. Each of the lines connects output from 7 different simulations which have equivalent C × t exposure products (e.g., a 0.5-h exposure at 400 ppm, a 4-h exposure at 50 ppm, and an 8-h exposure at 25 ppm). The PBPK model used was that published by Ward et al. (1988). Parameter values used are available elsewhere (Jarabek and McDougal, 1993).

Figure III-17 illustrates model simulations of different dose metrics of inhaled DCM at equivalent C × t exposure products of 200 ppm-hr. These different dose metrics would be appropriate to characterize different toxicities, depending on the choice of an assumed mechanism. For example, parent compound venous concentration (CV) and percent of carbon monoxide (CO) bound to hemoglobin (HbCO) could be chosen as the dose metrics for the neurotoxicity observed with DCM, because these effects have been attributed to both a non-specific narcotic action of the parent and to the hypoxic effect of its oxidative metabolic byproduct, CO (Winneke, 1981). Note that the profile for CV approximates a hyperbola on this plot, indicating that concentration is the major determinant (i.e., a plot of $C \times t^n \times t^m$ or $C \times t^{n+m}$ approximates the same shape as a plot of $C \times t^n$). For chronic toxicity, the amount of metabolite formed per gram liver tissue via the glutathione (GST) pathway might be

considered the appropriate metric since hepatic tumor incidence in mice has been shown to correlate well with the area under the curve for parent concentration in the liver and the amount metabolized via the GST pathway (Andersen et al., 1987b).

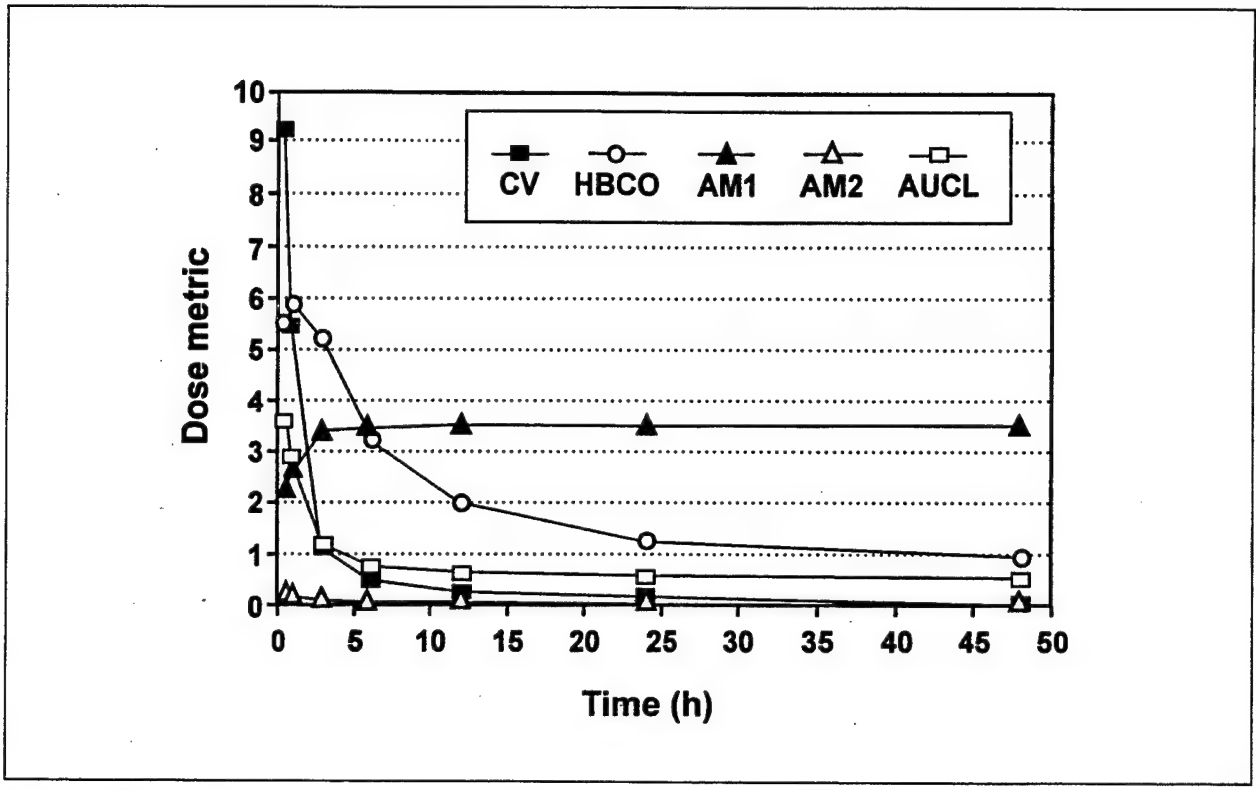


Figure III-17. Model Simulations of Different Dose Metrics in the Rat of Inhaled DCM at Equivalent C × t Exposure Products of 200 ppm-h. The PBPK model used was that published by Andersen et al., (1991). Parameter values used are available elsewhere (Jarabek and McDougal, 1993). CV = venous parent concentration (mg/L); HbCO = the percent of carbon monoxide bound to hemoglobin (%), AM1 = the amount of metabolite formed per gram liver tissue via the mixed function oxygenase system (mg-hr/g); AM2 = the amount of metabolite formed per gram liver tissue via the glutathione system (mg-h/g); and AUCL = the area under the curve for parent compound concentration in the liver (mg/L-h).

Although Figures III-15 through 17 illustrate simulations of equivalent C × t exposure products only for the rat, dosimetry models could be used to simulate the temporal profile of different dose metrics (e.g., those in Table III-5) for interspecies extrapolation. The model is exercised according to the experimental and objective scenarios for the laboratory animal species of interest (e.g., intermittent exposure regimen for rats and continuous exposure for humans), and the human equivalent concentration for a given observed effect in the laboratory animal is estimated as the exposure

concentration that results in an equivalent intensity of a chosen dose metric to that achieved with the experimental animal exposure from which the observed toxicity is extrapolated (U.S. EPA, 1994). Dosimetry model templates can be developed using default physiologic parameters (e.g., minute volume, blood flows) for the common laboratory animal species and humans. Chemical-specific physicochemical parameters (e.g., partition coefficients and metabolism rates) can then be used in these default templates. General categories for solubility of gases based on ranges of air:water partition coefficients (e.g., > 500, 10 – 500, < 10) could be used to develop models. A gas categorization scheme based on reactivity and water solubility has been used to generate default model structures (U.S. EPA, 1994; Jarabek, 1995). Limiting conditions for interspecies extrapolation can be defined by exercising the models to simulate extremes of key parameters and for different dose metrics. For example, models could be exercised to estimate exposures that result in equivalent parent and metabolite dose metrics (e.g., CV, AUBC, AUTC) between rat and human simulations for the extremes of high and low blood:air partition coefficient with high and low metabolic rates as bounds. O'Flaherty (1989) presented a similar framework with which to organize consideration of appropriate measures of delivered dose. Interspecies conversion of kinetically equivalent doses was proposed, based on systematic species dependencies of simple kinetic relationships between administered and delivered doses.

Because dosimetry models incorporate concentration and time dependent processes (e.g., rate of metabolism), time is explicitly accounted for and the default adjustment based on "Haber's Law" is obviated. These models also allow for development of interspecies relationships for different dose metrics. These dose metrics can be chosen on the basis of plausible mechanisms of action. The use of dosimetry models may therefore also provide revised definitions for "acute" versus "chronic" toxicity that take into account the dynamics of chemical disposition and damage. These models could also be used to simulate toxicity due to intermediate "less than lifetime" and intermittent exposures.

SUMMARY

Various environmental and regulatory statutes require risk characterization for exposure scenarios that range in duration from a few minutes to lifetime. Developing a dose-response estimate for such scenarios requires the use of available acute, subacute, subchronic, and chronic toxicity data and often the use of extrapolation procedures to different durations. The basis of current duration extrapolation procedures on "Haber's Law" and its attendant assumptions are presented. Toxicity depends on the magnitude, duration, and frequency of exposure. Choice of the appropriate dose metric and duration

extrapolation should depend on the mechanism of toxicity. Dosimetry models integrate mechanistic and temporal determinants of the exposure-dose–response continuum. Analysis of the limiting conditions for different mechanisms and dose metrics by chemical class categories is suggested as a promising approach to development of alternative extrapolation procedures.

ACKNOWLEDGMENTS

The author wishes to acknowledge collaborative work with Lt. Col. James N. McDougal, Ph.D. at the Toxicology Division, Armstrong Laboratory, Wright-Patterson Air Force Base, OH. Part of this work was presented as a poster at the 32nd annual meeting of the Society of Toxicology in New Orleans, LA, and is the subject of a manuscript in preparation.

REFERENCES

Andersen, M.E., MacNaughton, M.G., Clewell, H.J., III, and Paustenbach, D.J. 1987a. Adjusting exposure limits for long and short exposure periods using a physiological pharmacokinetic model. *Am. Ind. Hyg. Assoc. J.* 48(4):335–343.

Andersen, M.E., Clewell, H.J., III, Gargas, M.L., Smith, F.A., and Reitz, R.H. 1987b. Physiologically-based pharmacokinetics and the risk assessment process for methylene chloride. *Toxicol. Appl. Pharmacol.* 87:185–205.

Andersen, M.E., Clewell, H.J., III, Gargas, M.L., MacNaughton, M.G., Reitz, R.H., Nolan, R.J., and McKenna, M.J. 1991. Physiologically based pharmacokinetic modeling with dichloromethane, its metabolite, carbon monoxide, and blood carboxyhemoglobin in rats and humans. *Toxicol. Appl. Pharmacol.* 108:14–27.

Andersen, M.E., Krishnan, K., Conolly, R.B., and McClellan, R.O. 1992. Mechanistic toxicology research and biologically-based modeling: partners for improving quantitative risk assessments. *CIIT Activities* 12 (1):1–7.

Bliss, C.I. 1940. The relation between exposure time, concentration and toxicity in experiments on insecticides. *Ann. Entomol. Soc. Am.* 33:721–766.

Bliss, C.I., and James, A.T. 1966. Fitting the rectangular hyperbola. *Biometrics,* 22:573–602.

Corley, R.A., and Reitz, R.H. 1990. Dose-route extrapolation in quantitative toxicology: Physiologically based pharmacokinetics and pharmacodynamics of chloroform. In: Gerrity, T.R. and Henry, C.J., eds. *Principles of route-to-route extrapolation for risk assessment, Proceeding of the workshops;* March and July; Hilton Head, SC and Durham, NC. Elsevier Science Publishing Co., Inc. New York, NY. pp. 195–216.

Haber, F. 1924. *Zur geshichte des gaskrieges.* In: *Fünf Vorträge aus den Jahren* 1920–1923. Julius Springer, Berlin. pp. 76–92.

Hayes, W.J. 1975. Toxicology of Pesticides. The Williams and Wilkins Company, Baltimore. pp. 51–105.

Jarabek, A.M. 1994. Inhalation RfC methodology: Dosimetric adjustments and dose–response estimation of noncancer toxicity in the upper respiratory tract. *Inhal. Toxicol.*, 6:301–325.

Jarabek, A.M. 1995. The application of dosimetry models to identify key processes and parameters for default dose–response assessment approaches. *Submitted to Tox. Lett.*

Jarabek, A.M. and Segal, S.A. 1993. Noncancer toxicity of inhaled toxic air pollutants: Available approaches for risk assessment and risk management. In: D.R. Patrick, ed. *Toxic Air Pollutant Handbook.* Van Nostrand Reinhold, New York. pp. 100–127.

Jarabek, A.M. and McDougal, J.N. 1993. Duration-adjustment of effect levels: Comparison of physiologically-based pharmacokinetic (PBPK) models with default approach. *The Toxicologist* 13(1):282 (Abstract No. 1069).

O'Flaherty. 1989. Interspecies conversion of kinetically equivalent doses. *Risk Anal.* 9:587–598.

ten Berge, W.F., Zwart, A., and Appelman, L.M. 1986. Concentration-time mortality response relationship of irritant and systemically acting vapours and gases. *J. Hazard. Mater. 13:301–309.*

U.S. Environmental Protection Agency, 1993. [Reference concentration for chronic inhalation exposure (RfC) for HCFC-124 as verified 02/11/93.] Office of Health and Environmental Assessment, Environmental Criteria and Assessment Office, Research Triangle Park, NC.

U.S. Environmental Protection Agency. 1994. Methods for derivation of inhalation reference concentrations and application of inhalation dosimetry. Research Triangle Park, NC: Office of Health and Environmental Assessment, Environmental Criteria and Assessment Office; EPA report no. EPA/600/8-90/066F October 1994.

Ward, R.C., Travis, C.C., Hetrick, D.M., Andersen, M.E., and Gargas, M.L. 1988. Pharmacokinetics of tetrachloroethylene. *Toxicol. Appl. Pharmacol.* 93:108–117.

Warren, E. 1900. On the reaction of *Daphnia magna* (Straus) to certain changes in its environment. *Q. J. Microsc. Sci.* 43:199–224.

Winneke, G. 1981. The neurotoxicity of dichloromethane. *Neurobehav. Tox. and Terat.* 3:391–395.

ENDNOTES

a The NAS recommended that the scientific aspects of risk assessment should be explicitly separated from the policy aspects of risk management. Risk assessment was defined as the characterization of the potential adverse health effects of exposures to environmental hazards, and consists of four steps: (1) hazard identification: the determination of whether a chemical is or is not causally linked to a particular health effect; (2) dose–response assessment: the estimation of the relation between the magnitude of exposure and the occurrence of the health effects in question; (3) exposure assessment: the determination of the extent of human exposure; and (4) risk characterization: the description of the nature and often the magnitude of human risk, including attendant uncertainty.

b Apparently the only statement Haber made of what was to be called his rule is contained in a footnote to the last of a series of five lectures that this chemist made during the period 1920 through 1923 (Haber, 1924). The lecture pertained to the history of gas warfare and only brief exposures were considered. Also, at that time, no chemical was known that would not drift away or be diluted to a harmless concentration soon after its release (Hayes, 1975). The concept was actually not original with Haber but was stated first by Warren (1900) in connection with his studies of the effects of different concentrations of sodium chloride on *Daphnia magna*.

c Although the term physiologic based pharmacokinetic (PBPK) modeling is often used in a general sense, dosimetry modeling is used in this paper as a more comprehensive term to capture not only model structures used to address volatile organic chemicals but also irritant gases and particles. Mathematical modeling is defined as the use of the physical laws of mass, heat, and momentum conservation to quantify the dynamics of a system of interest (e.g., particle deposition and clearance in the respiratory tract). Dosimetry modeling is defined as the application of mathematical modeling to characterize the determinants of exposure-dose–response.

SESSION IV

EXPOSURE ASSESSMENT AND THE LIFE CYCLE TIMELINE

APPLICATION OF CULTURABLE SAMPLING METHODS FOR THE ASSESSMENT OF WORKPLACE CONCENTRATIONS OF BIOAEROSOLS

Martinez K.F.[1], Seitz T.A.[1], Lonon M.K.[2], and, Weber A.M.[1]

[1]National Institute for Occupational Safety and Health, Division of Surveillance Hazard Evaluations, and Field Studies, 4676 Columbia Parkway, MS-R11, Cincinnati, Ohio, 45226

[2] National Institute for Occupational Safety and Health, Division of Physical Sciences and Engineering 4676 Columbia Parkway, MS-R7, Cincinnati, Ohio, 45226

ABSTRACT

Case studies are presented demonstrating the utility of culturable air sampling methods as exposure assessment tools. These investigations included: (1) plants that manufacture enzymes; (2) a paper mill; (2) and a large office building with a ventilation system contaminated with *Penicillium.* In the first case study, a comparison of total bacterial counts (in combination with identification and quantification of the production strain) from unit processes to background locations identified exposure sites. Additionally, a comparison of the sampling results across the three manufacturing plants (among similar processes) identified effective control strategies based on the containment capabilities of the various technologies. This evaluative framework was also successfully applied in the second and third case studies. In combination with the identification and quantification of suspect microorganisms, emission patterns were identified to known immunologically active agents. In the second case study, elevated levels of *Thermoactinomyces* species were documented in the transfer tower and the biomass storage building. In the third case study, *Penicillium* from a contaminated ventilation system was identified as the predominant fungus in the indoor air. Current sampling methodologies for microbiological agents in ambient air are limited in their ability to comprehensively characterize personal exposures. However, these air sampling methods provide information that can be used to propose theories concerning agent dissemination and effectiveness of exposure control methods. In addition, when combined with medical and epidemiologic evidence, the collected data can help to establish causal relationships between exposures and symptoms.

INTRODUCTION

In the past, microbiologic exposure assessments focused primarily on the pathogenicity of microorganisms. We now know that certain microorganisms can produce metabolic by-products that can initiate toxemic reactions in exposed persons. The latter exposures are predominantly encountered in agricultural environments, such as mycotoxins produced by certain fungal species (e.g., aflatoxin from *Aspergillus flavus* and *Aspergillus parasiticus* species).[a,b,c] More recently, attention has been focused on the potential for microorganisms to promote immunologic responses in susceptible individuals. Immunologic responses are activated by an individual's challenge to particular antigenic constituents of a given microbial species which may be viable or non-viable. Allergic respiratory diseases resulting from exposures to microbiologic agents have been documented in agricultural, biotechnology, office, and home environments.[d,e,f,g,h,i,j,k] These responses and the subsequent expression of allergic disease is based, in part, on a genetic predisposition.[l] Allergic diseases typically associated with exposures in indoor environments include allergic rhinitis (nasal allergy), allergic asthma, allergic bronchopulmonary aspergillosis (ABPA), and extrinsic allergic alveolitis (hypersensitivity pneumonitis). The potential for immunologic effect and the emergence of new exposure environments has increased the need to elevate our understanding of microbial dissemination and exposure. Bioaerosol characterization is essential to the understanding and control of human exposures to these microbiologic agents.

Microorganisms (including fungi and bacteria) are ubiquitous inhabitants of the environment. Many of these organisms are saprophytic varieties (those utilizing non-living organic matter as a food source) that may inhabit soil, vegetation, water, or any reservoir that can provide an adequate supply of nutrient substrates. Under the appropriate conditions (optimum temperature, pH, and with sufficient moisture and available nutrients) microbial populations can proliferate. Through various mechanisms, these organisms can be disseminated into the environment as individual cells, cell aggregates, or in association with soil or dust particles or water droplets. In the outdoor environment, the levels of microbial aerosols will vary according to geographic location, climatic conditions, and surrounding activity. In indoor environments, the airborne concentration of microorganisms may vary somewhat as a function of the heating, ventilating, and air-conditioning (HVAC) system condition and the numbers and activity level of the occupants. Generally, indoor fungal levels in office environments are expected to be below the outdoor levels (depending on HVAC system filter efficiency) with consistently similar ranking among the microbial taxa.[m,n]

Acceptable levels of airborne microorganisms have not been established, primarily due to the varying immunologic sensitivities of individuals. Relationships between health effects and environmental microorganisms must be determined through the combined contributions of medical, epidemiologic, and environmental evaluations. The current strategy employed by the National Institute for Occupational Safety and Health (NIOSH) for on-site environmental evaluation involves a comprehensive inspection of the problem location and adjoining areas to identify sources of microbial contamination and routes of dissemination. In locations where contamination is visibly evident or suspected, bulk samples may be collected to identify the predominant taxa (fungi, bacteria, and thermoactinomycetes). In limited situations, air samples for microorganisms may be collected to document the airborne presence of a suspected microbial contaminant. Airborne dissemination (characterized by elevated levels in the complaint area, compared to outdoor and non-complaint areas, and anomalous ranking among the microbial taxa) which correlates with occupant symptomatology may suggest that the contaminant is responsible for the health effects.

Current air sampling methodologies for ambient microbiologic agents are limited in their ability to comprehensively characterize personal exposures. However, these air sampling methods provide information that can be used to propose theories on agent dissemination and effectiveness of exposure control methods. In addition, the collected data can, when combined with medical and epidemiologic evidence, help to establish causal relationships between exposures and symptomatology. Case studies are presented below demonstrating the utility of culturable air sampling methods as exposure assessment tools. These investigations included: (1) enzyme manufacturing processes; (2) a paper mill; (3) and a large office building with a ventilation system contaminated with *Penicillium*.

CASE 1: Enzyme Manufacturing Process

The advent of recombinant DNA technology and the subsequent ability to engineer novel microorganisms posed a variety of concerns about the adequacy of conventional biotechnology processes to contain the microbiologic production strains. In addition to occupational exposures to viable microorganisms, this technology may also result in exposures to biologically active products or intermediates and to processing chemicals such as extraction solvents. This case study describes the results of an investigation to evaluate the effectiveness of controls in enzyme fermentation processes.[P] The control measures were evaluated by collecting air samples for substances involved in enzyme manufacturing operations.

Approximately 200 air samples for bacteria were collected at each of three enzyme manufacturing plants. Air samples were collected with Andersen two-stage viable cascade impactors at a calibrated flow rate of 28.3 liters per minute (lpm). The 50% effective cutoff diameter for the Andersen sampler is 8 μm — hence, larger, non-respirable (not lung depositable) particles are collected on the top stage and smaller, respirable (lung depositable) particles are collected on the bottom stage.[q,r] Sampling locations were selected to estimate workplace exposures to production microorganisms from process equipment. These locations included a microbiology laboratory, areas near inoculum and fermentor tanks, locations contiguous to filtering operations, and various outside "background" locations. For each facility, sampling was conducted over a five-day period with sample times ranging from 2½ to 20 minutes. Standard Methods Agar was used as the sampling culture medium for the enumeration of total bacteria. Identification of the production strain was facilitated through colony morphology, gram stain, and/or sugar utilization profile.

Microbial concentrations of bacteria varied from plant to plant due to the microorganisms used in production and the type of process equipment utilized (results are summarized in Figure IV-1). At one plant, levels of total culturable airborne bacteria at a solid-liquid separation process (filter press) averaged 6820 colony forming units per cubic meter of air (CFU/m^3), the predominant strain being the production microorganism. At a second plant, levels of culturable bacteria at a separation process (rotary vacuum drum filter) equipped with a local exhaust ventilation hood averaged 345 CFU/m^3, but the production strain was identified in very low numbers on only 10% of the sample plates. At a third plant, levels of culturable bacteria at a separation process (centrifuge) averaged 875 CFU/m^3, with the production strain present in significant numbers. The airborne bacterial levels around the rotary vacuum drum filter at Plant 1 were considerably lower than the levels around the filter press at Plant 2 presumably due to the inherently better containment characteristics of rotary vacuum drum filters (low energy) and the use of local exhaust ventilation. Based on the inherent design of centrifuges (process isolation) at Plant 3, microbiological emissions would be expected to be high compared to filter presses. However, due to process enclosure and the use of local exhaust ventilation, airborne microbiologic levels (total and production microorganism) were elevated at the centrifuge, but at significantly lower levels than the filter press. Airborne bacteria levels in the laboratory, at the inoculum tank, and at the fermentor tank were all below the levels obtained at the filter operations.

Based on the results of air sampling, the greatest potential for occupational exposures to production bacterial strains exists around high energy operations including filters or centrifuges, agitator shafts,

and manual fermentor sampling ports. These operations often are not amenable to complete sealing, enclosure, or isolation. However, where total enclosure of a potential emission source is not a feasible alternative, local exhaust ventilation is shown to be an effective means in controlling microbial emission sources (as applied in Plant 1 at the rotary vacuum drum filter and in Plant 3 at the filter press), thereby reducing the potential for exposure. In addition, the work practices of the operators can be a determining factor in the degree of exposure. For example, the bacterial concentrations observed during filter press operations (Plant 1) were primarily the result of worker interaction with the process. Additionally, the improper practices of workers during the extraction of broth samples from the fermentor tank sample ports resulted in aerosolization of microbial residues in the piping system.

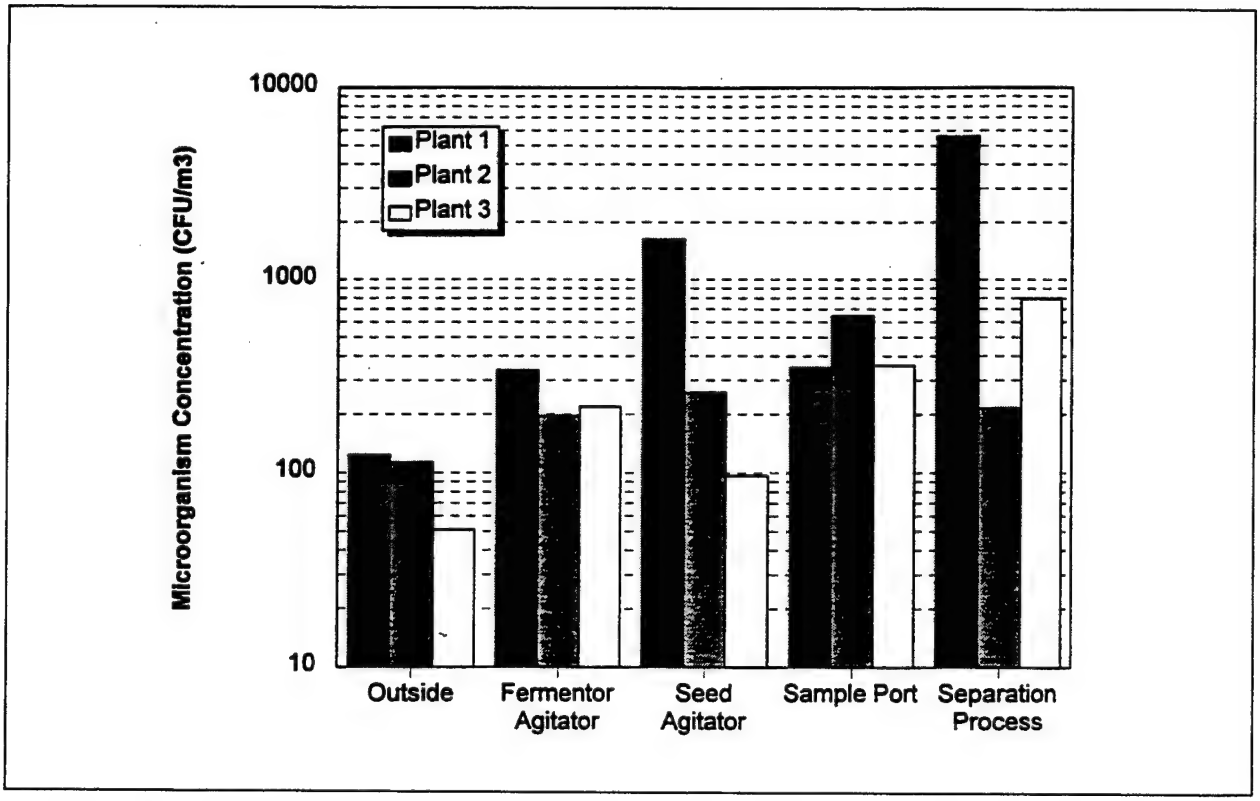


Figure IV-1. Airborne Bacterial Concentrations at Select Processes in Enzyme Manufacturing Operations.

CASE 2: Paper Mill

The most common microorganisms involved in the development of hypersensitivity pneumonitis are the bacterial species (*Micropolyspora faeni*, *Thermoactinomyces vulgaris*, *Thermoactinomyces saccharin*, and *Thermoactinomyces candidus*). Various members of the class *Thermoactinomyces* (TA) have been implicated in the development of the hypersensitivity disease "farmers lung," as a result of their presence in moldy fodder. Hypersensitivity to wood dust, pulp, and chips has been described in the literature.[s,t,u,v] The manufacture of paper products requires the storage of large stockpiles of wood logs. These stockpiles are kept moist and provide an environment conducive to the growth of microorganisms, especially thermotolerant species. In response to a health hazard evaluation request concerned with two employee fatalities and other employee health concerns, investigators with NIOSH conducted a study at a paper mill to evaluate chemical and microbiological exposures.[w]

In order to determine the concentrations of airborne microorganisms at selected process locations, the Andersen 2-stage viable cascade impactor was used at a calibrated flow rate of 28.3 lpm. Trypticase Soy Agar was used for the enumeration of bacteria. The sample plates for bacteria were incubated at 50 °C to promote the growth of thermotolerant bacteria (*Thermoactinomyces*). Sampling times varied according to the estimated load on the sample plates at different process locations. Sampling locations included the transfer tower (where wood chips and treatment plant biomass conveyor systems converge), the A-frame building (used for biomass storage), the de-barking building, the pulp room, and outdoors (for background reference). Temperature and relative humidity were recorded for each sample run at each sampling location.

Observation of Figure IV-2 indicates that the highest concentrations occurred in the transfer tower and the A-frame building (the left axis of the graph is referenced to the vertical bars describing bacterial concentration and the right axis of the graph is referenced to the lines describing the respirable and TA fractions). These concentrations were two orders of magnitude greater than those observed in the outdoor samples. Additionally, TAs from these sampling locations comprised greater than 90% of the genera, whereas, the percentage of TAs in the outdoor samples comprised only 45% of the total bacterial population sampled. The numbers of thermotolerant bacteria in the de-barking area and the pulp room do not appear to be significantly different from the outdoor samples. However, the percentages of TAs are higher in these locations compared to the samples collected outdoors. Notably, the respirable fraction observed in the transfer tower was significantly higher than all other locations (approaching 80%).

The concentration of TAs in the transfer tower and the A-frame building (as compared with the concentrations in the outdoor samples) support the conclusion that the bark of the wood used for the paper making process is a source of these microorganisms. The absence of similar concentrations in the de-barking area can be directly attributed to the application of water to the wood logs during the bark removal operation. This "wetting" action reduces the potential for bacteria entrained dust or wood particulates to become airborne. Due to the elevated presence of TAs in the transfer tower and the A-frame building, recommendations were made to reduce employee exposures. Recommendations included limiting access to essential personnel, the application of engineering controls, and medical surveillance for cases of hypersensitivity pneumonitis in exposed personnel.

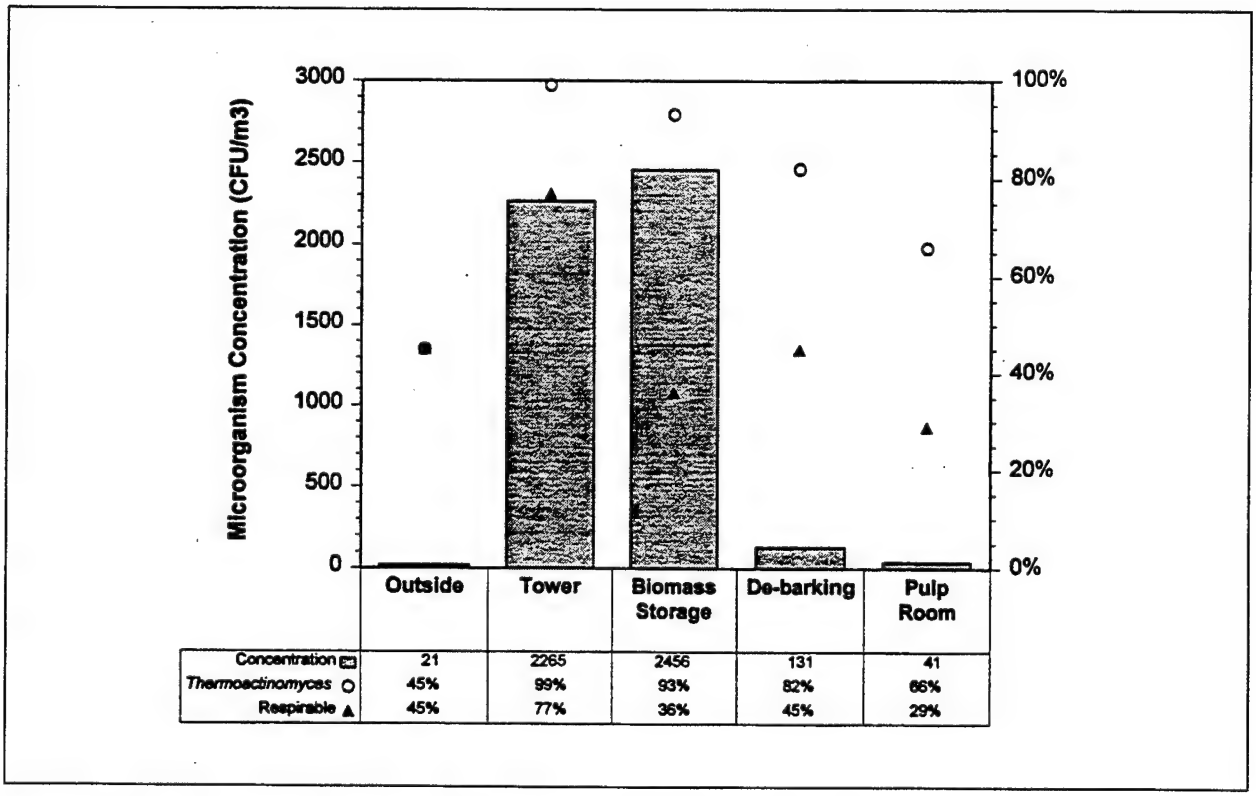


Figure IV-2. Airborne Bacterial Concentrations at a Paper Mill.

CASE 3: Large Office Building with a Contaminated HVAC System

The relationship of microbiological contamination to the onset of symptoms considered characteristic of indoor environmental quality problems (specifically, Sick Building Syndrome) is not fully understood. Furthermore, cases of medically-diagnosable Building Related Disease (hypersensitivity pneumonitis, occupational asthma, and allergic rhinitis) do not always correlate with environmental sampling data. Effective sampling protocols for bioaerosols require an *a priori* awareness of microbial dissemination patterns, sampling methods that can indicate "true" immunologic challenge, and adequate replication of samples to confidently estimate exposures. To gain a better understanding of the relationship between microbiological contamination and occupant immunologic effects and symptoms, a cross-sectional morbidity study was initiated in a large office building. Localized "mildew" odors were traced to microbial contamination in a ventilation system employing fan coil units (FCUs) to treat occupied zones. Inherent design flaws in the ventilation system (improper cooling coil water temperature, low efficiency air filters, and porous duct lining) had permitted the development, amplification, and dissemination of microbial reservoirs (predominantly *Penicillium*).[x]

Approximately 156 air samples for culturable airborne fungi (10 minute sampling times with the Andersen 2-stage viable cascade impactor at a calibrated flow rate of 28.3 lpm) were collected at 12 interior building locations and one outdoor location (on the roof). Twelve replicate samples for each location were collected on two separate days over a three day period. Attempts were made to randomly allocate sample teams to sample locations on different days to minimize bias due to across day effects and sampling team effects. Temperature and relative humidity were recorded for each sample replicate. Characterization of microbial contaminant (culturable fungi) dissemination from the HVAC system diffusers was facilitated by the Spiral Air Systems (SAS) Compact Portable Air Sampler (Pool Bioanalysis Italiana, Milano, Italy). All culturable samples were collected over a sampling time of 2.7 minutes at a calibrated flow rate of 109 lpm. All samples were collected approximately 8 centimeters from the diffuser. Approximately five replicates were collected at each sample location. Malt Extract Agar was used for the enumeration of fungi on all sample plates (Andersen and SAS) incubated at room temperature. Morphological characteristics were used to determine the genera of fungi and their rank orders were noted.

Data analysis was facilitated by grouping sampling locations of similar clusters. Grouped sampling locations were classified as Roof, Floor 2, Floor 3, Floor 4, and Complaint. The complaint area was defined by the nature and number of occupant symptoms which prompted the investigation. A

graphical summary of the sampling data (geometric means) for culturable fungi as collected with the Andersen impactors is presented in Figure IV-3. The left axis of the graph is referenced to the vertical bars describing fungal concentration (CFU/m^3); the right axis of the graph is referenced to the lines describing the respirable and *Penicillium* fractions. An analysis of variance (ANOVA) was used to compare logarithmically transformed group sampling location means. Statistically significant differences were observed (Bonferonni corrected significance level, = 0.01) between the roof samples and all other indoor sampling locations (p-value < 0.001). Additionally, statistically significant differences (= 0.01) were observed between the complaint area versus all other indoor locations (Floor 2, Floor 3, and Floor 4) and between Floor 2 versus Floor 3 (p-value < 0.001). No significant differences (= 0.01) were observed between Floor 2 versus Floor 4 (p-value = 0.048).

In the complaint environment, 95% of the fungal genera identified were *Penicillium*. In contrast, the roof and Floor 3 air samples exhibited mean *Penicillium* percentages of 13 and 36%, respectively. The mean *Penicillium* percentage on Floor 2 was 77% and on Floor 4 (excluding Section 5) the mean was 45%. The variation in the rank order of predominant genera between the outdoor and indoor environments indicates the presence of a microbiologic reservoir and microbiologic dissemination, especially in the complaint area. The respirable fraction closely correlates with the percentages of *Penicillium* for all sample locations except outside. This is consistent with the expected size range of *Penicillium* spores (2 to 6 μm), which would be expected to impact on the lower stage of the Andersen sampler (50% effective cut-off diameter at 8 μm).

A graphical summary of the culturable fungi sample results from the SAS sampler is presented in Figure IV-4. ANOVA was used to compare group sampling location means. Statistically significant differences were observed (significance level, = 0.05) between samples collected in the complaint environment and other indoor locations (Floor 3 and Floor 4) and all other indoor sampling locations (p-value < 0.001). In Section 5, 96% of the genera identified were *Penicillium*. In contrast, Floor 3 and Floor 4 samples were predominated by 42 and 58% *Penicillium*, respectively. Because all SAS samples were collected at the supply air diffuser, high percentages of *Penicillium* on sample plates indicate *Penicillium* spore dissemination from contaminated ventilation systems. The results from the SAS sampler correlated well with the culturable fungi results from the Andersen samplers. For example, in the complaint environment, the Andersen samplers indicated a geometric mean concentration of 129 CFU/m^3 with 94% of the fungal isolates identified as *Penicillium*, while the SAS sampler indicated a geometric mean concentration of 142 CFU/m^3 with 96% of the fungi identified as

Penicillium. Sampling results showed significantly higher concentrations of airborne, culturable *Penicillium* in the complaint environment (in the occupied spaces and from the supply air diffusers) than in all other building areas sampled. This data documents the dissemination of *Penicillium* to the occupied spaces from reservoirs in the FCU systems (particularly in the complaint environment).

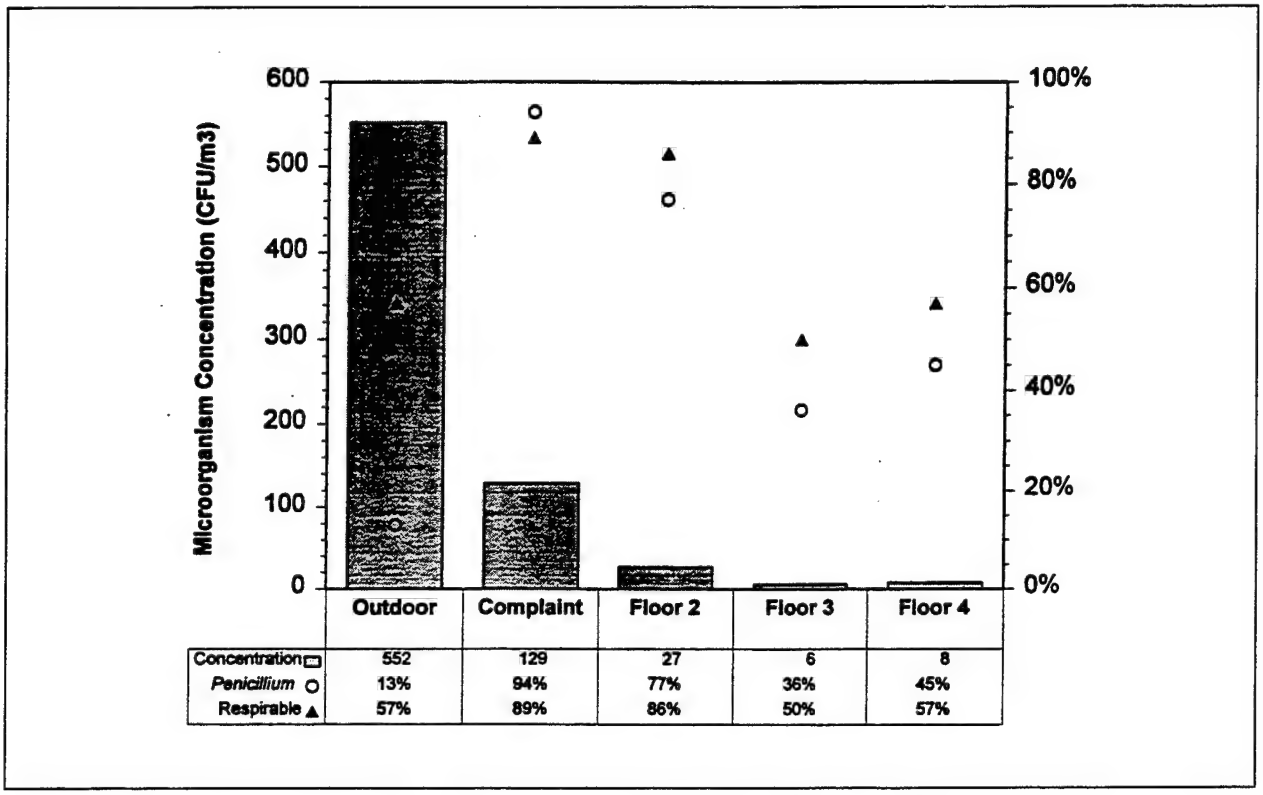


Figure IV-3. Airborne Fungal Concentrations in a Large Office Building.

Although many of the building FCU systems had some level of microbial contamination, the unique environment in the complaint area appeared to promote the amplification of microbiological reservoirs (specifically *Penicillium*). The moisture level in the FCUs has a significant influence on the degree of microbiologic amplification. Given the relative humidities encountered in the occupied spaces (approaching or exceeding 60%), the humidity levels in the area past the cooling coils of the FCUs would be expected to range up to 95%, as observed during the 1992 sampling survey.

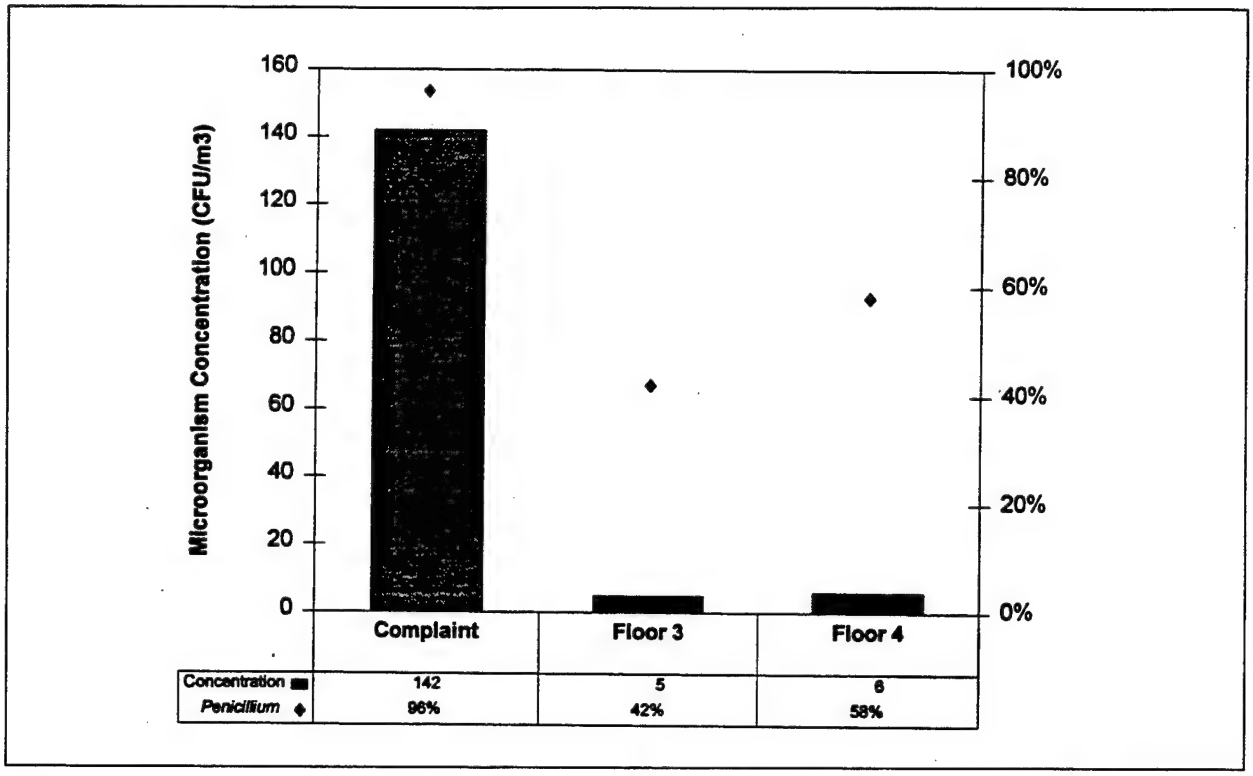


Figure IV-4. Airborne Fungal Concentrations at Supply Air Diffusers (SAS sampler).

CONCLUSIONS

In the first case study, culturable air sampling methods were used to document workplace concentrations of bioaerosols around enzyme manufacturing process equipment. A comparison of total bacterial counts (in combination with identification and quantification of the production strain) from unit processes to background locations identified exposure sites. The microbiologic emissions from process equipment primarily resulted from equipment containment deficiencies and/or poor operator work practices. Additionally, a comparison of the sampling results across the three manufacturing plants (among similar processes) identified effective control strategies based on the containment capabilities of the various technologies.

This evaluative framework was also successfully applied in the second and third case studies. Replicate samples at each location allowed comparisons between evaluated areas and background locations. In combination with the identification and quantification of suspect microorganisms, emission patterns were identified to known immunologically active agents. In the second case study, elevated levels of *Thermoactinomyces* species were documented in the transfer tower and the A-frame

building. In the third case study, *Penicillium* from a contaminated ventilation system was identified as the predominant fungus in the indoor air.

The use of culturable air sampling techniques is shown to be an effective means of assessing bioaerosol concentrations in various occupational environments. However, a holistic evaluation with environmental and medical components is needed to determine the relationship, if any, between bioaerosol exposures and worker symptoms or illnesses. It should be noted that immunologic reactions in susceptible individuals can also be activated by non-viable biologic products or components as well as by viable microorganisms. Therefore, comprehensive investigations of microbiologic exposures should include non-culturable air sampling techniques, where appropriate, in order to estimate the contribution of non-viable particulates. Non-viable sampling methods primarily involve the use of collection media that facilitate microscopic analyses of the samples, such as membrane filters and glass slides (i.e., spore traps).[y,z,aa,bb]

Non-viable biological elements and non-culturable organisms, as well as culturable microorganisms, may be responsible for infection and toxicosis, in addition to immunologic responses. There is a need for alternative methods of sampling and analysis that focus on the composition and quantity of specific biochemical analytes. For instance, recent studies report the use of ergosterol, a component of fungal membranes, as an indicator of total fungal biomass.[cc,dd,ee] Rylander et al. describes an evaluation of air quality in Swedish "sick" buildings through the collection of airborne particulates onto filters with subsequent analysis for -1-3-glucans (fungal cell wall constituents).[ff] Analysis of bacterial polysaccharides and fatty acids may be used for the total and differential determination of bacteria in aqueous solutions.[gg] In cases where a specific microbiologic agent is the primary focus, analysis of filter samples subjected to polymerase chain reaction (PCR) analysis shows promise for a wide variety of applications. Sawyer et al. report the successful application of PCR analysis to track the dissemination of Varicella-Zoster virus from the individual rooms of diagnosed patients to other areas of the hospital.[hh] Although the efficacy of many of these methods for the analysis of air samples has yet to be determined, future applications of the tools of analytical chemistry and molecular biology will supplement and complement the methods currently used and will contribute to a better understanding of the microbial ecology of indoor air and its potential health effects on exposed individuals.

REFERENCES

Banazak, E.F., Barboriak, J., Fink, J., Scanlon, G., Schlueter, E.P., Sosman, A., Thiede, W., and Unger, G. [1974]. Epidemiologic studies relating thermophilic fungi and hypersensitivity lung syndrome. *American Review of Respiratory Disease.* **110**:585–591.

Blomquist, G., Palmgren, U., and Ström, G.. [1984]. Improved techniques for sampling airborne fungal particles in highly contaminated environments. *Scandinavian Journal of Work Environmental Health.* **10**:253–258.

Burg, W.R., Shotwell, O.L., and Saltzman, B.E. [1981]. Measurements of airborne aflatoxin during the handling of contaminated corn. *American Industrial Hygiene Association Journal.* **42**(1):1–11.

Burge, H.A. [1988]. Environmental allergy: definition, causes, control. *Engineering Solutions to Indoor Air Problems.* Atlanta, GA: American Society of Heating, Refrigeration and Air-Conditioning Engineers pp. 3–9.

Buttner, M.P. and Stetzenbach, L.D. [1993]. Monitoring airborne fungal spores in an experimental indoor environment to evaluate sampling methods and the effects of human activity on air sampling. *Applied and Environmental Microbiology. 59*(1):219–226.

Commercial Biotechnology: An International Analysis [1984]. Washington D.C.. U.S. Congress, Office of Technology Assessment. OTA-BA-218.

Curtis, S.E., Balsbaugh, R.K., and Drummond, J.G. [1977]. Comparison of Andersen eight-stage and two-stage viable air samplers. *Applied and Environmental Microbiology.* **35**(1):208–209.

Eduard, W., Lacey, J., Karlsson, K., Palmgren, U., Ström, G., and Blomquist, G. [1990]. Evaluation of methods for enumerating microorganisms in filter samples from highly contaminated occupational environments. *American Industrial Hygiene Association Journal.* **51**(8):427–436.

Edwards, J.H. [1980]. Microbial and immunological investigations and remedial action after an outbreak of humidifier fever. *British Journal of Industrial Medicine.* **37**:55–62.

Fink, J.N., Banaszak, E.F., Thiede, W.H., and Barboriak, J.J. [1971]. Interstitial pneumonitis due to hypersensitivity to an organism contaminating a heating system. *Annals of Internal Medicine.* **74**:80–83.

Fox, A., Rosario, R.M.T., and Larsson, L. [1993]. Monitoring of bacterial sugars and hydroxy fatty acids in dust from air conditioners by gas chromatography-mass spectrometry. *Applied and Environmental Microbiology.* **59**(12):4354–4360.

Gardner, R.M., Tindall, G.W., Cline, S.M., and Brown, K.L. [1993]. Ergosterol determination in activated sludge and its application as a biochemical marker for monitoring fungal biomass. *Journal of Microbiological Methods.* **17**:49–60.

Gessner, M.O., Bauchrowitz, M.A., and Escautier, M. [1991]. Extraction and quantification of ergosterol as a measure of fungal biomass in leaf litter. *Microbial Ecology.* **22**:285–291.

Grant, W.D. and West A.W. [1986]. Measurement of ergosterol, diaminopimelic acid and glucosamine in soil: evaluation as indicators of microbial biomass. *Journal of Microbiological Methods.* **6**:47–53.

Hodgson, M.J., Morey, P.R., Attfield, M., Sorenson, W., Fink, J.N., Rhodes, W.W., and Visvesvara, G.S. [1985]. Pulmonary disease associated with cafeteria flooding. *Archives of Environmental Health.* **40**(2):96–101.

Jäppinen, P., Haahtela, T., and Liira, J. [1987] Chip pile workers and mould exposure: a preliminary clinical and hygienic survey. *Allergy.* **42**:545–548.

Malmberg, P., Rask-Andersen, A., Palmgren, U., Höglund, S., Kolmodin-Hedman, B., and Stålenheim, G. [1985]. Exposure to microorganisms, febrile and airway-obstructive symptoms, immune status and lung function of swedish farmers. *Scandinavian Journal of Work and Environmental Health.* **11**:287–293.

Martinez, K.F., Sheehy, J.W., and Jones, J.H. [1988]. Control technology assessment of enzyme fermentation processes. Cincinnati, OH: *National Institute for Occupational Safety and Health.* Publication No. 88–114.

Martinez, K.F. and Wilcox, T.G. [1994]. *NIOSH health hazard evaluation interim report:* Social Security Administration Building, Baltimore, Maryland. HETA 92-0355.

Morey, M.R. and Feeley, J.C. [1990]. The landlord, tenant, and investigator: their needs, concerns and viewpoints. *Biological Contaminants in Indoor Environments.* Baltimore, MD: American Society for Testing and Materials pp. 1–20.

Nevalainen, A., Willike, K., Liebhaber, F., Pastuszka, J., Burge, H., and Henningson, E. [1993]. Bioaerosol sampling. *Aerosol Measurement: Principles, Techniques, and Applications.* Willike K., Baron P., editors. 471–492.

Pickering, C.A. [1992]. Immune respiratory disease associated with the inadequate control of indoor air quality. *Indoor Environment.* **1**:157–161.

Rylander, R., Persson, K., Goto, H., Yuasa, K., and Tanaka, S. [1992]. Airborne -1-3-glucan may be related to symptoms in sick buildings. *Indoor Environment.* **1**:263–267.

Sawyer, M., Wu, Y., and Wallace, M. [1994]. Detection of varicella-zoster DNA in air samples from hospital rooms. *The Journal of Infectious Diseases.* **169**:91–94.

Schlueter, D.P., Fink, J.N., and Hensley, G.T. [1972]. Wood-pulp workers' disease: A hypersensitivity pneumonitis caused by . *Annals of Internal Medicine.* **77**:907–914.

Seitz, T.A., Parker, J.E., Tubbs, R.L. and Martinez, K.F. [1993]. *NIOSH health hazard evaluation report: Champion International*, Bucksport, Maine. HETA 90-375-2334.

Sorensen, W.G. [1989]. Health impact of mycotoxins in the home and workplace: an overview. *Biodetermination, Degradation, Mycotoxins, Biotoxins, and Wood Decay.* O'rear CE, Llewellyn GC, editors. Plenum Press. pp. 201–215.

Sosman, A.J. [1969]. Hypersensitivity to wood dust. *New England Journal of Medicine.* **281**(18):977–980.

Sorensen, W.G., Simpson, J.P., Peach, M.J., Thedell, T.D., and Olenchock, S.A. [1981]. Aflatoxin in respirable corn dust particles. *Journal of Toxicology and Environmental Health.* **7**(3):669–672.

Terho, E.O., Husman, K., Kotimaa, M., and Smoblom, T. [1980]. Extrinsic allergic alveolitis in a sawmill worker. *Scandinavian Journal of Work Environmental Health.* **6**(2):153–157.

Topping, M.D., Scarsbrick, D.A., Luczynska, C.M., Clarke, E.C., and Seaton, A. [1985]. Clinical and immunological reactions to among workers at a biotechnology plant. *British Journal of Industrial Medicine.* **42**:312–318.

Turner, A.G. and Hill N.F. [1975]. Calibration of the Andersen 2000 disposable air sampler. *American Industrial Hygiene Association Journal.* **36**(6):447–451.

Vinken W. And Roels P. [1984]. Hypersensitivity pneumonitis to in compost. *Thorax.* **39**:74–74.

Weiss, N.S. and Soleymani, Y. [1971]. Hypersensitivity lung disease caused by contamination of an air-conditioning system. *Annals of Allergy.* **29**:154–156.

THE ROLE OF HISTOPATHOLOGY OF THE TESTIS IN SHORT-TERM TOXICOLOGY TESTING PROTOCOLS

Lonnie D. Russell

Laboratory of Structural Biology, Department of Physiology
Southern Illinois University School of Medicine, Carbondale, IL 62901-6512

ABSTRACT

A re-examination of methodology employed for histopathological evaluation of the testis is currently underway in light of changing and upgraded regulatory requirements/guidelines. With the current general trend toward shortening of testing protocols it will be imperative to change the way the testis is examined. Although testis cells are exposed to a toxicant, usually only some will be affected and the effect seen upon histological examination is dependent on a variety of factors which are considered herein. Histopathology must be more sensitive than in the past for the detection of treatment-related effects. To improve the sensitivity of this endpoint, it is necessary to obtain better quality tissue for the histopathologist to examine and to have histopathologists well-trained in evaluation of the testis. Quantitation of testis germ cell types is fundamentally important to provide objective information for risk assessment analysis, either to show that the tissue is normal or to show the degree to which spermatogenesis is depressed.

INTRODUCTION

Historically, evaluation of the reproductive systems in the male and female for toxic effects has assumed a low priority. Interference with one's ability to reproduce was considered less important than interference with other bodily functions. However, the development of assisted reproductive techniques such as *in vitro* fertilization and the cadre of related techniques have underscored the need for many couples who are infertile to pursue medical means to salvage fertility. The media's role in bringing reproductive problems and the technology to solve them to the forefront of our attention was paramount.

With public consciousness focused on reproduction competence as a medical issue, several examples of adverse toxic effects on reproduction have also served to heighten public awareness that the reproductive systems should be considered in environmental, occupational and pharmaceutical

exposure to agents. Regulatory agencies as well as the private sector have had to face the question of how to test for potential reproductive harm. The multigeneration study, organ weights (sometimes obtained) and traditional histopathology have been the methodological mainstay by which differences in exposed and control groups have been evaluated. Soon it was realized that fertility determinations in rodents were insensitive (Blazak et al., 1985) and that histopathology was in need of upgrading to obtain the maximum amount of information from the tissue. New guidelines have been proposed/formulated by many regulatory agencies. Most recently, there has been an attempt to standardize testing protocols necessary prior to human exposure to certain pharmaceutical agents. The trend in such discussions has been to shorten protocols, and in many instances, protocols have been upgraded, requiring all to re-examine what endpoints are the most important prior to human exposure[a].

Histopathology is the most useful parameter in reproductive evaluation. What follows is an examination of the impact of the potential changes in regulatory guidelines/requirements on histopathological evaluation of the testis, with emphasis given to those relating to shortening of testing protocols. Specifically, this paper addresses current and anticipated problems with histopathological evaluation of the testis and the ways in which these can be solved.

THE PROBLEM(S)

The value of current endpoints to assess reproductive toxicity is severely compromised by protocols that call for shortening of the time in which the animal is exposed to an agent. To understand the impact of shortened exposure protocols, it is first necessary to also understand the timing of the spermatogenic process and the timing of sperm movement through the epididymis, keeping in mind how and when agents may act during this process.

THE TIMING AND VULNERABILITY OF SPERMATOGENESIS

Spermatogenesis is a process whose timing is thought to be constant for a particular strain of animal but slight differences in the timing for various strains have been recorded and major species differences occur (Russell, et al., 1990). Although various estimates have been made for the length of spermatogenesis, in the Sprague-Dawley rat, the entire process of spermatogenesis lasts about 53 days. Epididymal transit of sperm takes approximately 1.5 weeks, bringing the time for the total process to about 63 days or nine weeks (Figure IV-5). Early spermatogenesis, including the most immature of the germ cells, the spermatogonia, has been shown to be vulnerable to many agents that effect the genome or the cell cycle. It must be assumed that according to extensive literature on the topic accumulated

over many years, agent exposure has the potential to affect the process at any point. Epididymal effects are less commonly seen than those in the testis but can not be dismissed from consideration.

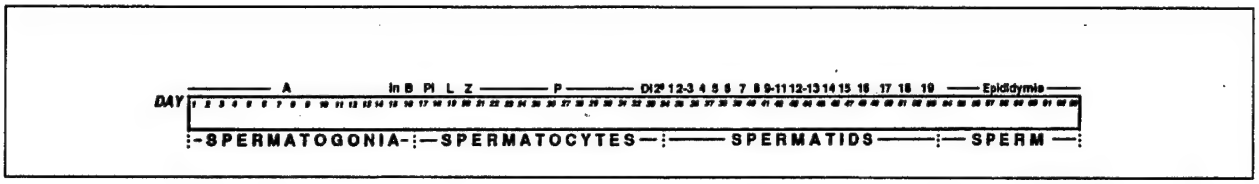


Figure IV-5. The Time Course of Spermatogenesis is Illustrated as a Horizontal Bar with the Days Numbered in the Upper Portion of the Bar. The diagram also shows the germ cell types present at these times and the period of time that sperm reside in the epididymis. Abbreviations are used in this and subsequent figures to denote germ cell types present at specific times during spermatogenesis. (A= Type A spermatogonia; In = intermediate spermatogonia; TypeB = B spermatogonia; Pl = preleptotene spermatocytes; L= leptotene spermatocytes; Z = zygotene spermatocytes; P = pachytene spermatocytes; Di = diplotene spermatocytes; 2^0 = secondary spermatocytes; large numbers from 1 through 19 are the steps of spermiogenesis and indicate spermatid development; small numbers inside the boxes indicate days that correspond with developmental phases of spermatogenesis.

The Temporal Impact of Experimental Design on Histopathological Effects in the Male

Although the suggested period of exposure to agents differs widely from regulatory agency to regulatory agency[a], assume that the exposure period to an agent is of the shortest duration now being discussed. Exposures recommended for the male are as short as four weeks (repeated exposure) with some indication that a two-week exposure might be acceptable. A two-week agent exposure and a fertility trial of two weeks duration is used herein. It illustrates the greatest impact of a short-term protocol on spermatogenesis. Figure IV-5 is the template used to illustrate the course of spermatogenesis. It is used to show where and when agent exposure could potentially effect and impact the spermatogenesis and epididymal sperm transit. This map portrays the 63-day process on a horizontal scale with abbreviations of the cell types present at particular periods of the process to show the cell types developing during that period.

Various scenarios should be considered to understand the impact of short-duration exposure periods on spermatogenesis. The complete list of assumptions that are necessary to show how and when agents can impact spermatogenesis are detailed in the legend to the pertinent Figures. The histopathological picture noted, given that the most mature germ cells were affected (Figure IV-6X) would appear as early as day 1 after immediate sacrifice and if the most immature cells were affected (Figure IV-6Y) as late as day 63. These two scenarios marked 6X and 6Y in Figure IV-6 carry the

assumption that, due to a bioaccumulation phenomena, it was necessary to reach an agent threshold in the two weeks exposure to produce an affect on the most immature and mature germ cells present. Only one cell type was affected at one brief period in spermatogenesis. In the former case (Figure IV-6X) it was a rarely seen Type A stem cell spermatogonia that would have most likely not been detected in any histopathological examination. In the latter scenario (Figure IV-6Y) the alteration in the effected cells (sperm) would not be detected either by histopathology and probably would not be detected by fertility tests.

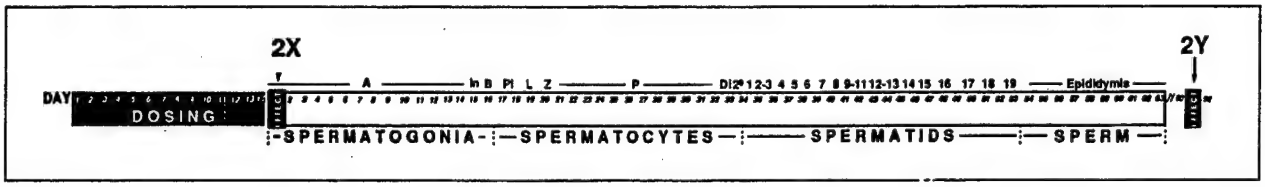


Figure IV-6. The Horizontal Bar is Used in Figure IV-5 is Utilized to Show Two Scenarios (X and Y) of How an Exposure(s) Can Effect Early and Late Cells in the Spermatogenic Process. The assumptions made in each of these scenarios are as follows:

Scenario X

Exposure Duration: Two weeks
Sacrifice Interval (time between end of exposure and sacrifice): Immediate
Bioaccumulation Time Necessary (time to reach a toxic threshold): Two weeks
What was Affected by the Agent: One germ cell type at one stage at the very beginning of the cycle.
What was Seen by the Histopathologist: Effect on Type A spermatogonia at one stage.

Scenario Y

Exposure Duration: Two weeks
Sacrifice Interval: Immediate
Bioaccumulation Time Necessary: Two weeks
What was Affected by the Agent: Sperm cells at the very end of epididymal storage
What was Seen by the Histopathologist: Nothing

An infinite number of other possible scenarios are possible, but only a few will be illustrated. For example, Figure IV-7 shows an effect from exposure on the spermatocyte population of cells that was manifest immediately and lasted the duration of the exposure. Such a pattern indicates that only one cell type was affected during the course of exposure since the time span of the effect is of the same duration as the exposure. In Figure IV-8 the effect is immediate and lasted the duration of the exposure but, unlike the example provided in Figure IV-7, it is broader in that it affected more than one cell type.

In Figure IV-9 and 10 effects are produced on one cell type at one specific developmental phase in an experiment that has zero (Figure IV-9) and one that has a three week (Figure IV-10) time lapse

between exposure and histopathological examination. The effect can be picked up by histopathology in the example illustrated in Figure IV-9, but is not detected by either histopathology or fertility tests in Figure IV-10 because the effected cells have matured and moved from the male reproductive tract.

In some instances one may determine when during the spermatogenic process that cell type was exposed/affected from a diagram like that shown in Figure IV-10 or from available software that has been specifically designed for this purpose (Hess and Chen, 1992). Determinations of this type can be made if the paradigm for agent effects follows the example shown in Figure IV-7 in that the histopathological effect is of the same duration as the exposure and the effect is considered to occur simultaneously with the administration and the effect is considered to end at the time that the exposure ends. Such an effect like that illustrated in Figure IV-7 creates what has been called a window of effect when it is examined histopathologically. In other words there is an effect, whether it be loss of germ cells or abnormal germ cells, that occurs between normal appearing germ cells that are younger and others that are older.

The window of effect is only seen when the effect is clearly noted to begin at a certain point in the spermatogenic process and end likewise and, as such, can be easily related to maps such as in Figure IV-7. But not all affects on the testis have clear beginnings and endings. There may be permanent damage to the testis due to an agent effect. Such effects may by caused by agent effects to stem cell spermatogonia or to somatic cells such as Sertoli cells although the cause for most permanent effects is not readily ascertained. If Sertoli cells are injured, it is generally, but not always the case, that the effect will be manifest on several germ cell types simultaneously and that the effects will be scattered and not as readily be indicated on staging maps (Anderson, et al., 1989).

Assuming a window of effect is evident, the longer the sacrifice interval the more the chance the window of effect has had to progress. Figure IV-11 shows a window of effect moving through spermatogenesis as a function of sacrifice interval. Only when the window of effect reaches the epididymal sperm (Figure IV-11Y) can fertility be potentially affected. It is obvious from such a diagram that fertility may be unaffected by some treatments that have previously affected spermatogenesis (Figure IV-11Z). In such instances and instances where only the testis is affected (Figure IV-11W–X), the only way to ascertain damage is to utilize histopathology.

From the foregoing scenarios it is easy to see that sort-term protocols will result in a variety of responses. Most are related to the testis since insufficient time has transpired to allow the window of effect to reach the epididymis where final maturation of sperm occurs and functional sperm are stored.

Consider an alternative type of protocol where dosing occurs for nine weeks and mating trials take place during the last two weeks of exposure. Even employing assumptions where bioaccumulation phenomena took as long as four weeks for an effect to reveal itself by histopathology, the results would appear like that shown in Figure IV-12 . Detection of such an effect would be less difficult because the window of effect would be much broader. Thus, the foregoing scenarios show that shorter term protocols place more of a burden on the histopathologist to find the effect if it is present. Evaluation of short-term protocols has also been published by Harris et al. (1992).

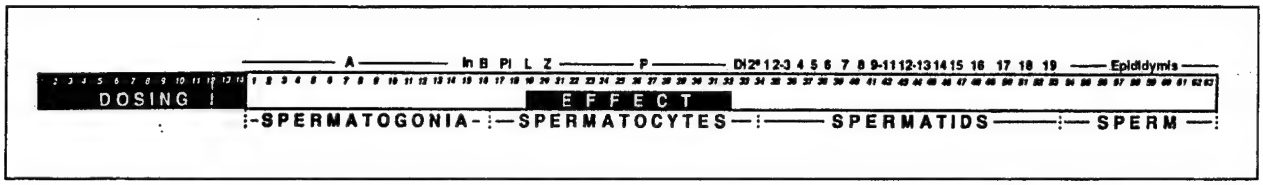


Figure IV-7. The Horizontal Bar from Figure IV-5 is Utilized to Show How an Effect Could Span the Same Period of Time in Terms of Cell Developmental as the Period of the Exposure.

Exposure Duration: Two weeks

Sacrifice Interval: None

Bioaccumulation Time Necessary: none

What was Affected by the Agent: One arbitrarily selected spermatocyte type at one stage.

What was Seen by the Histopathologist: A two-week duration of effect in the spermatocyte population

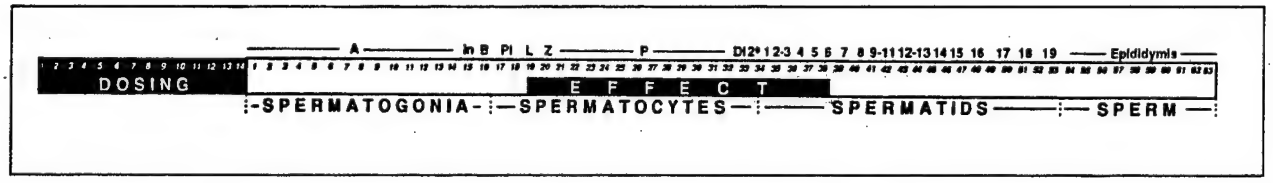


Figure IV-8. The Horizontal Bar from Figure IV-5 is Used to Illustrate How an Effect Could Be Broader Than the Period of Exposure.

Exposure Duration: Two weeks

Sacrifice Interval: None

Bioaccumulation Time Necessary: none

What was Affected by the Agent: More than one arbitrarily selected cell type in the spermatocyte population of cells was affected by the exposure.

What was Seen by the Histopathologist: A 20-day developmental effect in the spermatocytes and spermatids.

Figure IV-9

Figue IV-10

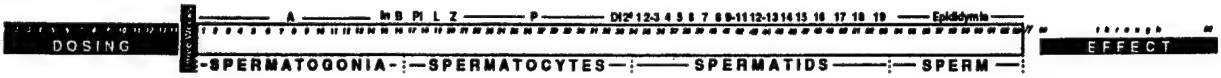


Figure IV-9 and 10. The Horizontal Bar from Figure IV-5 is Used to Illustrate How Sacrifice Interval Effects the Cells Affected during the Spermatogenic Process. In Figure IV-9 there is no sacrifice interval and in Figure IV-10, there is a three-week sacrifice interval.

(Figure IV-9)

Exposure Duration: Two weeks

Sacrifice Interval: None

Bioaccumulation Time Necessary: none

What was Affected by the Agent: One arbitrarily selected cell type in the spermatid population of cells was affected by the exposure.

What was Seen by the Histopathologist: A two week effect in the spermatids

(Figure IV-10)

Exposure Duration: Two weeks

Sacrifice Interval: Four weeks

Bioaccumulation Time Necessary: none

What was Affected by the Agent: One cell type in the spermatocyte population of cells (not indicated) was affected.

What was Seen by the Histopathologist: No effects

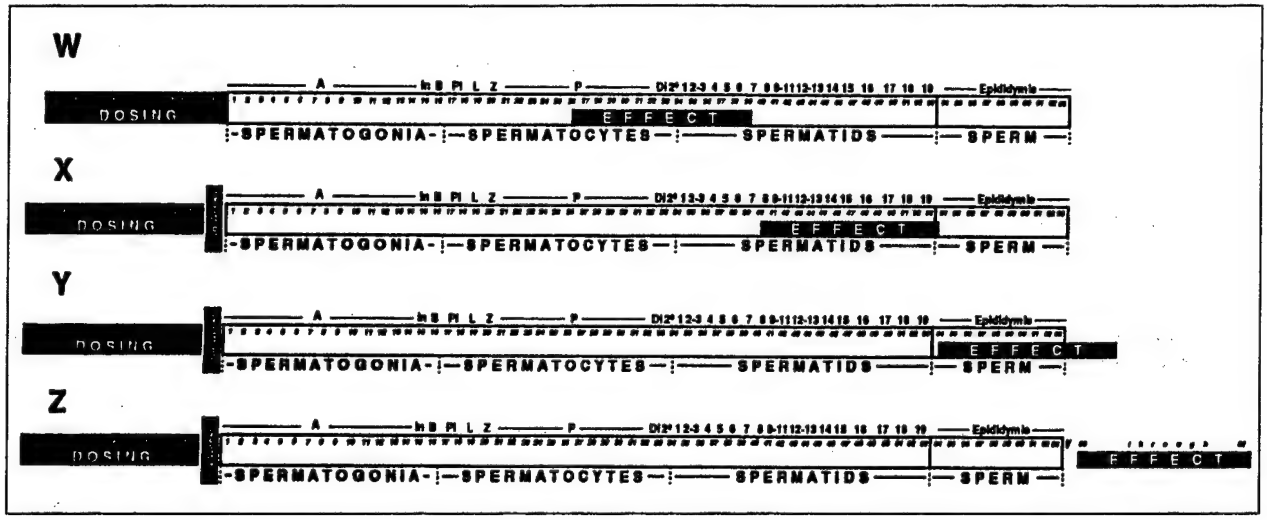


Figure IV-11. The Horizontal Bar from Figure IV-5 is Used to Illustrate the "Movement" of the Effected Cells through the Spermatogenic Process as a Function of Sacrifice Interval.

Exposure Duration: Two weeks

Sacrifice Interval: None- Figure IV-11W; Two weeks Figure IV-11X; Four weeks Figure IV-11Y; Six Weeks Figure IV-11Z

Bioaccumulation Time Necessary: none

What was Affected by the Agent: One cell type at one stage of the spermatocyte population of cells was affected by the exposure.

What was Seen by the Histopathologist: for Figure IV-11W effects that are seen in spermatocytes & spermatids. For 11X effects are seen in spermatids. For 11Y no effects are seen in the testis. For Figure IV-11Z no effects are noted histologically

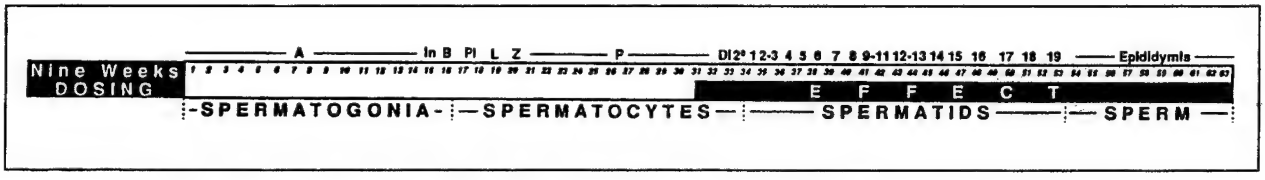


Figure IV-12. The Horizontal Bar from Figure IV-5 is used to Illustrate the Effect Seen in an Experiment of Long Duration in Spite of Immediate Sacrifice and in Spite of Lengthy Bioaccumulation Time. Although in previous figures both bioacculation time and immediate sacrifice were detrimental to noticing an effect utilizing histopathology, the long duration of the experiment overcame such problems.

Exposure Duration: Nine weeks

Sacrifice Interval: None

Bioaccumulation Time Necessary: Four weeks

What was Affected by the Agent: One cell type at one stage of the spermatocyte population of cells was affected by the exposure.

What was Seen by the Histopathologist: Effects were noted in spermatocytes and spermatids.

CURRENT HISTOPATHOLOGY

Histopathologists must be competent to evaluate numerous tissues of the body and have been trained to do so quite well. The enormity of their overall task should not be under-emphasized nor should their past and continuing contributions to the risk assessment data gathering process. However, the reality of the situation is that histopathologists are generally poorly trained in testis evaluation. This is not a reflection on the skill of histopathologists but a comment about the complexities of testis structure and about the state of our knowledge of testis pathology which is relatively new compared with most other organ systems. Furthermore, evaluation of the testis is especially complex. Although a recent book focuses specifically on upgrading this endpoint (Russell et al., 1990), the average histopathologist commonly detects only major differences between the control and treatment groups. Damage detection is of enormous value given that, even if difference are not specifically classified, they are, at least, noted and can be used to some degree in risk assessment prediction.

Unfortunately, most protocols do not allow histopathologists to follow spermatogenesis as a process to determine differences in normal and agent-treated groups. To understand spermatogenesis as a process requires good tissue preparation and a understanding of staging (see below).

Currently, tissues are processed using Bouin's, or more commonly, formalin fixation. They are embedded in paraffin and stained with hematoxylin and eosin or PAS. These techniques allow mainly identification of nuclei with little specific staining of cell constituents.

Histopathology, from a historical standpoint, is designed to evaluate the qualitative appearance of the testis. At best, most histopathologists use only semi-quantitative techniques to demonstrate germ cell loss. As an example, the histopathologist devises and employs a semi-quantitative scheme to grade the severity of a response into five or so categories that are somewhat arbitrarily defined.

Histopathology, even if performed with the most sensitive eye and undertaken by a highly trained histopathologist, can not usually detect the small changes in cell populations that might occur after short-term treatments. This is the case especially, if those changes are in numbers of very immature germ cells, then the response may be almost impossible to detect on routine examination of the testis. Loss of a few spermatogonia at some stages of the cycle, as viewed in each tubular cross sections, will eventually lead to a complete absence of most advanced germ cells in the tubule Ð a profound effect. Given that there are just a few spermatogonia in most tubules, their loss is difficult to detect.

SOLVING THE PROBLEM(S)

Modern histopathology of the testis must accommodate the anticipated changes in testing protocols by improvements in the sensitivity of such exams. There are several ways in which this can be accomplished. Some are currently feasible and others require further development.

IMPROVED TISSUE PREPARATION

Suggestions that histopathological procedures be upgraded have been made in several evaluations of the topic (Russell, 1983; Chapin, 1988; Russell et al., 1990; Hess et al, 1993). Moreover, testing guidelines from several regulatory agencies suggest/mandate upgrading tissue preparation methodologies. These suggestions have taken the form of improvements both in methodology to prepare tissue and in the education of the histopathologist.

There is no general agreement on what is needed in histotechnique to provide adequate information on stained slides. It is well known that minor variations in technique such a using Bouin's fixative will allow better appearance of tissues than using formalin and indeed many programs have already switched to this fixation procedure.

It has been advocated that formalin fixation followed by methacrylate embedding will serve as an adequate upgrade (Chapin et al., 1984). Indeed, there is less shrinkage and cellular detail is improved. This technique represents a major improvement over that currently in general use.

The author's personal preference is to utilize Bouin's fixative followed by paraffin embedding for most animals in a treatment group, but for about one-quarter to one-sixth of the animals, employ perfusion fixation of animals with subsequent embeddment in epoxy. When viewing slides in the epoxy-embedded group, the histopathologist has the opportunity to obtain the maximum amount of information from the tissue in preparations of the kind that show exquisite detail and which highlight degenerating cells (Figure IV-13). Identification of cell degeneration is a key to finding pathology since there is a dictum known as "die or go", applying to all but undifferentiated spermatogonia, which indicates that cells will either continue to progress at their normal pace as they progress through spermatogenesis or they will die. Usually, dead cells remain conspicuous within the epithelium in epoxy-embedded and stained material for a day or so before they are degraded. The histopathologist, given sufficient information on the slides of tissue embedded in epoxy, can then examine slides fixed using the Bouin's technique. Although these slides show less detail it is relatively easy to verify changes seen before in perfused, epoxy-embedded tissues.

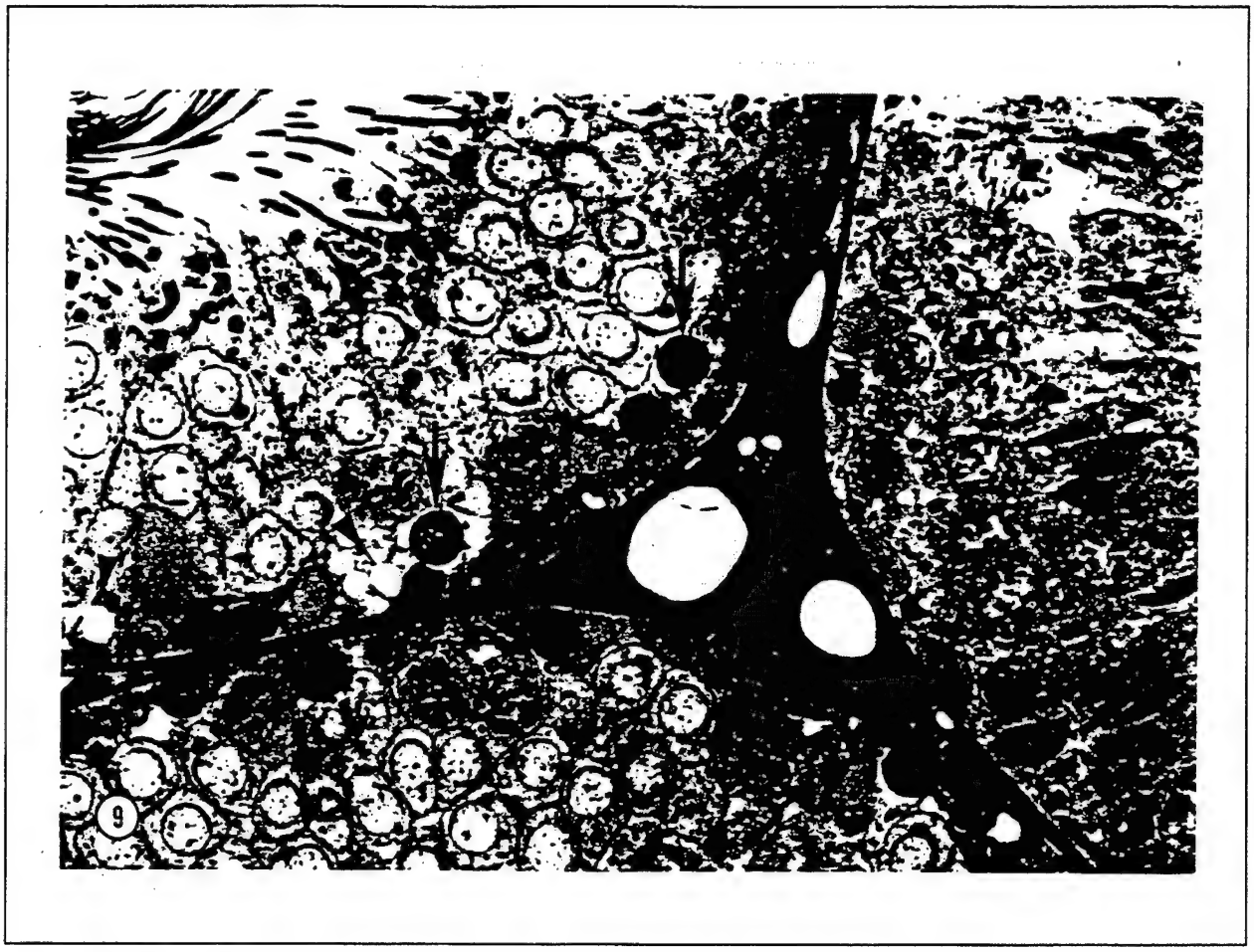


Figure IV-13. A Typical Example of the Appearance of Testis Tissue Fixed in Glutaraldehyde and Embedded in Epoxy. A few degenerating cells are indicated by arrows. Vesicles related to Sertoli cells are indicated by arrow heads.

UPGRADING THE HISTOPATHOLOGIST KNOWLEDGE OF THE TESTIS

Upgrading the histopathologists knowledge about histopathology and staging is an important and necessary requirement, allowing the histopathologist to take advantage of the information available from the tissue. This recommendation should be tempered by the enormous job already confronting the histopathologist, especially when histopathologists are required to evaluate all major tissues of the body and to examine hundreds of slides in one day.

This problem can be overcome by task specialization and/or peer review. In any group of histopathologists, one individual can undergo specialized training in testis and epididymal evaluation. Such training courses are available. That individual may be responsible for peer review of all male

reproductive tract slides and may be called upon to serve as a knowledge resource for others in the group.

UTILIZING STAGING

A knowledge of staging allows one to follow spermatogenesis as a process since tubules do not show each phase of the process. Actually, in a cross-sectioned tubule one only obtains a small glimpse of the entire process. To gain an elementary knowledge of staging takes one or three devoted days for one to become comfortable with the spermatogenic process. A book has been written (Russell, et al., 1990) and software is currently in the final stages of development (by Rex Hess) that will assist the investigator in learning staging. By using staging it is possible to predict what cell type will be affected from a knowledge of the cell types showing abnormalities.

Staging maps are constructed by breaking up diagrams like Figure IV-14X into short segments that are placed on top of each other (Figure IV-14Y). Using pictures of the germ cells instead of alphabetical symbols, a staging map can be constructed for species such as the dog (Figure IV-15). One follows the progress of spermatogenesis from left to right along a horizontal axis until the end of a horizontal row where the process is continued in the row above beginning at the left. That there are multiple horizontal rows in a staging map allows one to relate what is seen in a staging map to a cross-sectioned seminiferous tubule. Each vertical column designates what can be seen in a cross-sectioned seminiferous tubule as a cell association. Use of a staging map facilitates universal communication of effects from one scientist to another whether it be within an institution to another or from an institution to a regulatory agency.

A chart such as that shown in Figure IV-16 can be utilized to show how one can predict the cell type affected by a particular treatment from the histopathological evaluation. This can only be done when the cell type affected is affected immediately and no bioaccumulation phenomena occur.

QUANTITATION OF TESTIS GERM CELL TYPES

Quantitation of testis cells using absolute numbers lends itself to statistical analysis. Thus, one may show with a 95% probability whether a treatment does or does not cause an effect. In the latter situation (no difference between groups), the regulatory agency has much more confidence in the data than is provided by subjective histopathology. In the latter, the extent to which spermatogenesis is depressed can be quantified and the severity of the response in various dose groups compared. Risk

assessment is facilitated when a quantitative evaluation is undertaken and statistical analysis is supplied.

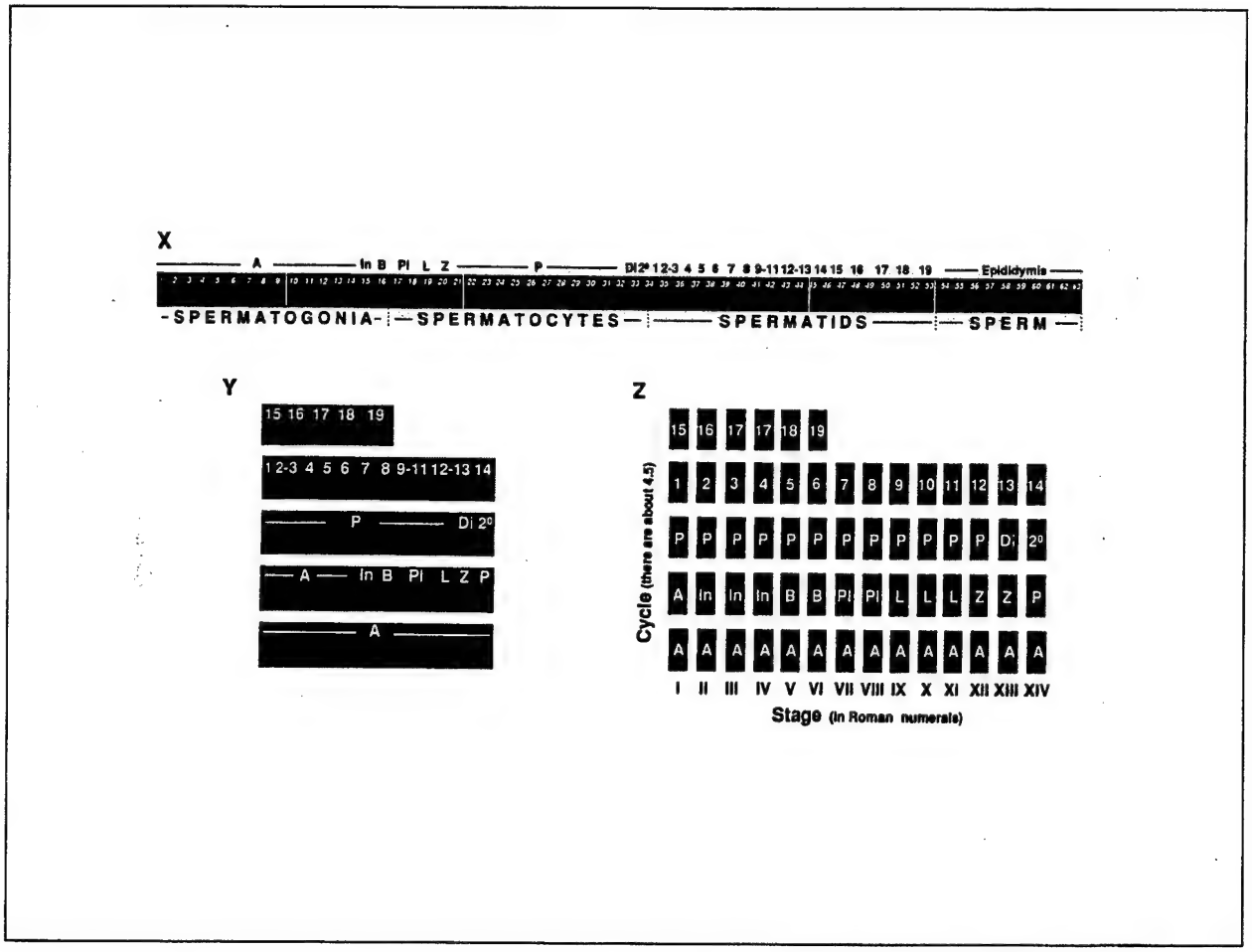


Figure IV-14. Construction of a Crude Staging Map (14Z) from the Horizontal Bar (14X) used in Previous Figures to Depict the Time Course of Spermatogenesis. In Figure IV-14X the bar is broken up into segments at the vertical lines and stacked one on another. Thus spermatogenesis is followed by going from the left to the right until the end of the bar and picking up again at the left of the bar immediately above. In 14Z the bars are broken into segments which represent cell types. What is seen in vertical rows are cell associations that represent what is seen in cross-sectioned tubules. Each cell association is given a Roman numeral. The horizontal rows are cycles, each lasting comparable periods of time. There are approximately four and one-half cycles for most species. The designations utilized represent cell types.

Methods for quantitation of testis germ cell types have been devised (Amann, 1970, Berndtson, 1977; Russell, et al., 1990). Most are research methodologies but a few have implications for obtaining quick answers in a toxicological setting where large numbers of animals are examined. The few that

may be applicable are described below. It is clear that the direction of future research should emphasize the development of methods that provide quick and objective quantitation to complement qualitative histopathological findings.

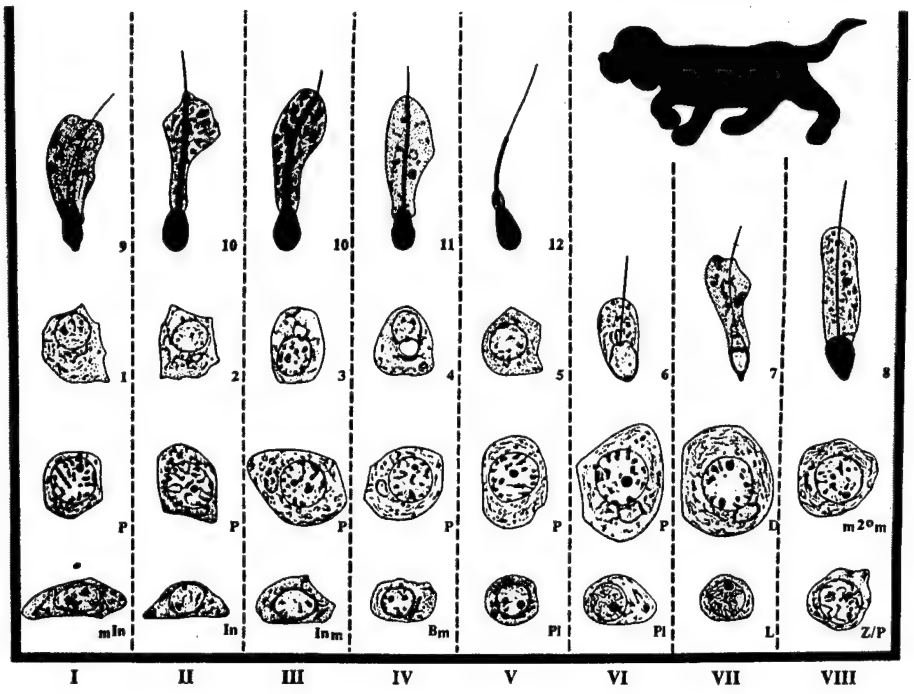


Figure IV-15. A Staging Map of the Dog Spermatogenesis Showing Illustrations of Cell Types Instead of Letter Designations used in Earlier Figures. (Used with permission from Histological and Histopathological Evaluation of the Testis published by Cache River Press, Clearwater, FL)

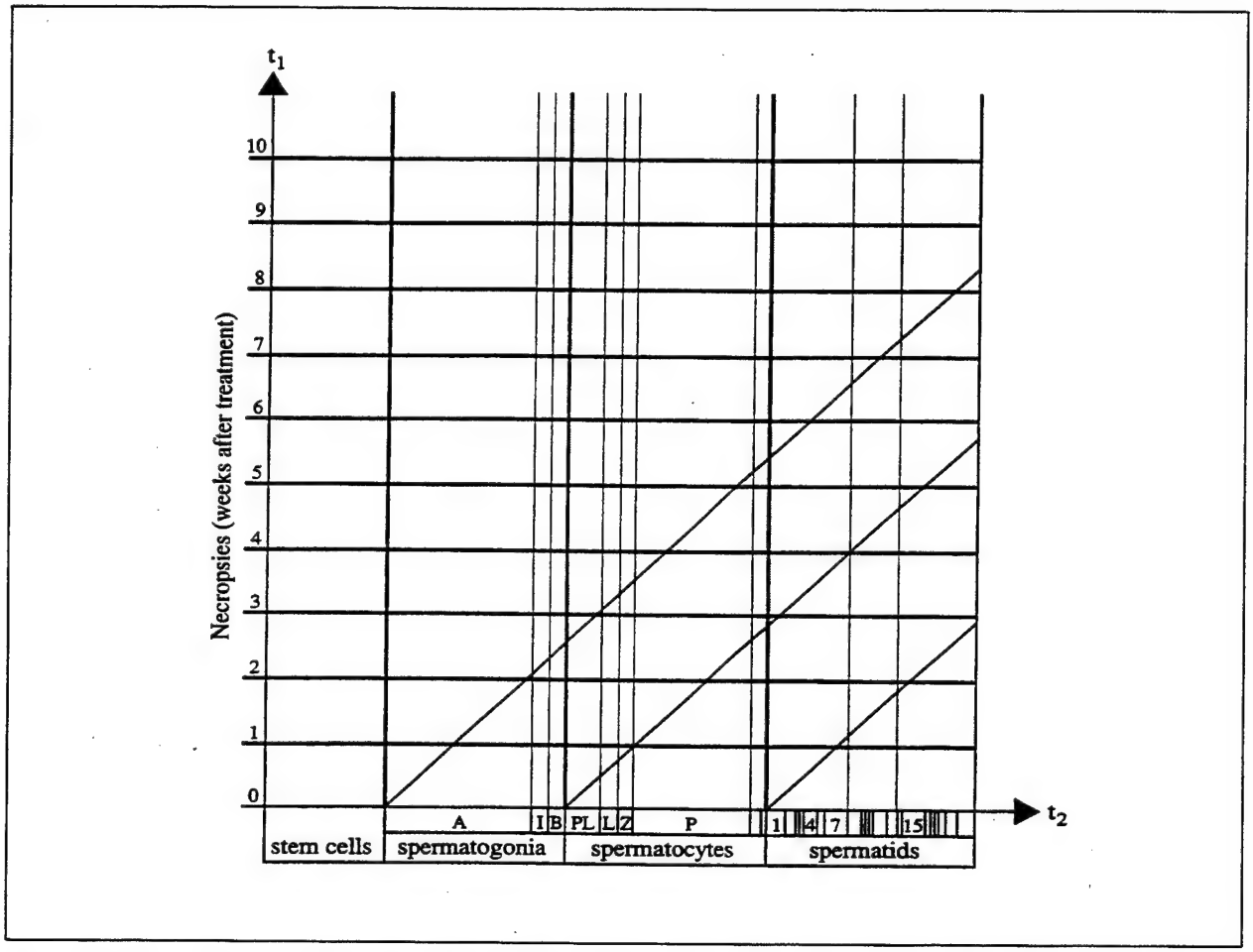


Figure IV-16. Chart used for Prediction of the Cell Affected from the Observance of an Effect by Histopathology. To do so accurately requires that the effect be manifest without bioaccumulation. It is also best to have a window of effect that is easily distinguished. To use this chart place a point at the intersection of the least advanced germ cell type affected and the sacrifice interval. Make a line (parallel to those already in place) from the point to the base of the chart to see where the line intersects the cell types listed (by abbreviations) to determine what cell type was affected. (From Ettlin et al., 1984 Archives of Toxicology used with permission).

TESTIS WEIGHT AND TUBULAR DIAMETER MEASUREMENTS AS INDICATORS OF TUBULAR CONTENTS

These techniques do not quantify germ cells. They have used for some time and have been good predictors that of quantitative changes in cell populations within seminiferous tubules. If there is a significant decrease in either of these parameters (or even a trend), it is even more incumbent on the individual to quantify tubular contents. Tubular diameters are measured in cross- or nearly cross-

sectioned tubules in well-fixed tissue. One should be aware of the large number of animals needed to detect differences in groups with an 80% chance of detecting a 10 or 15% difference (Berndtson, 1989) since tubules at different stages of the cycle have differing diameters (Wing and Christensen, 1982).

HOMOGENIZATION RESISTANT SPERMATID COUNTS TO PREDICT DAILY SPERM PRODUCTION

The end product of spermatogenesis are the sperm so it would seem natural to seek means to evaluate the production of testicular sperm. Once sperm are released in the process of spermiation they move down the duct system to the epididymis so fast that they are rarely seen within the tubules of the testis (Russell, 1993). Thus, to count the departing testicular sperm is not practical.

However, counts of almost-mature spermatids is possible using a technique of homogenization of testis cells with a blender and subsequent detergent digestion followed by hemocytometer counts of resistant sperm nuclei. Detergent digestion with Triton-X will cause all but the most packed (condensed) spermatid nuclei to lose their recognition in light microscope preparations. Counts of recognizable elongate spermatids can be plugged into an appropriate formula that considers the duration of stages that the nuclei are resistant to detergent for a particular species to determine the number of sperm produced per day (Amann, 1970).

This technique represents a relative rapid and inexpensive means to determine sperm production. It is most useful in experiments in which dosing has been continuous for an extended period. In such instances sperm production rates will have stabilized, albeit at a lower level than controls. However, in short-term experiments, a deficit in sperm production may be undetectable if it has not impacted the nearly-mature spermatid population of cells. Only when the window of response moves through the almost-mature population of elongate spermatids would a deficit be noted. There is also some question that, in addition to recording sperm heads from normal sperm, the technique will also record sperm that have not been released as the result of failure of sperm release (Russell et al., 1990; Ku et al., 1993).

There is disagreement as to the sensitivity and repeatability of homogenization techniques to measure sperm production. Some indicate that it is both sensitive and reproducible (Blazak et al, 1993). However, within a single laboratory, data from the bull and stallion indicate that to record a significant 10% (20%) decline in testis weight ($p < 0.05$), one must utilize 330 (83) and 1,350 (377) animals, respectively (Berndtson, 1989). Chapin et al found it the least sensitive of a number of endpoints (Chapin et al., 1992). Several studies have shown that counts of homogenization resistant

spermatids correlate poorly with fertility tests in rodents. Measurement of sperm production may yield very different results if data across labs are compared.

COUNTS OF SPERM RESERVES

A technique similar to that used to determine sperm production is used to count sperm stored in the epididymis. Sperm reserves are those sperm that are in the process of passing through the epididymis while they are maturing. Sperm reserves also reflect sperm that are stored in the cauda epididymis. It reflects past production of sperm as well as how full/empty the epididymis is at the time of evaluation. This rapid and low cost method (Robb et al., 1978) is the only rapid method to determine sperm reserves, but it is expected that, like in the determination of sperm production, the sensitivity of the test is low compared with a good evaluation of testis histopathology.

FLOW CYTOMETRY

Flow cytometry may be used to analyze "testicular" and epididymal sperm for detection of chromosomal abnormalities, changes in DNA content, and amount of single-stranded versus double-stranded DNA (Evenson et al., 1983). Counts of cells can also be made using flow cytometry.

Specifically tagged DNA in testis germ cells and somatic cells is detected with flow cytometry (Evenson et al., 1985). The data obtained expresses numbers of cells in three broad categories related to DNA. For example, all somatic cells (Leydig cells, Sertoli cells, myoid cells and peritubular cells among others) have a 2n (diploid) complement of chromosomes; most spermatogonia also have a 2n complement; almost all primary spermatocytes and G_2 spermatogonia are 4n (tetraploid) ; secondary spermatocytes are also 2n and spermatids are n (haploid).

Flow cytometry has been proposed as a tool to use in conjunction with histopathology to determine changes in cell populations and has been validated for toxicological studies (Evenson et al., 1985). The method appears to have high sensitivity but low specificity with respect to cell types measured. For example, all somatic cells and some germ cells are represented in a single count. Some increase in specificity is obtained if seminiferous tubules are first isolated before tissue is prepared for flow cytometry. This is a laborious procedure and can not be considered as practical in a toxicology testing protocol setting. Also, the purchase and maintenance of a flow cytometer dedicated to this process represents a major expense.

AUTOMATION OF GERM CELL COUNTS IN THE TESTIS

Ideally, counts obtained from tubules would satisfy the need for quantitation of cells in control and treatment groups and would allow statistical analysis. Manual counting of hundreds of cells within the epithelium is laborious, taking 30–45min/tubule. Such a task is prohibitive given the numbers of tubules needed to give low variation in statistical analysis (Berndtson, 1989).

As indicated above, risk assessment is facilitated when information can be analyzed statistically from objective measurements or, in this instance, cell counts. Computerized counts of the various types of testis cells at one or more stages of the spermatogenic cycle would seem to be an ideal means to show that the testis germ cell population is either statistically similar or is different from the treated tissue. There have been no published accounts where computers have been employed to count cells, although this goal seems feasible considering current image analysis technological advances and capabilities in movement and storage of images obtained from the light microscope. Development of this methodology would represent an important step to complement current histopathological evaluation of the testis.

In addition, it is feasible to develop computer software that could translate specific losses in testis components to see if they predict testis weight and, more importantly, if they could determine the effect of decrements in cell counts on the eventual production of sperm.

Combination Studies

Protocols have been designed to obtain numerous measures of reproductive capability using the same animals. Animals undergoing toxicology testing have blood, two testes, two epididymides and reproductive function. Careful design of protocols should allow the use of virtually all of these parameters for a more comprehensive view of the male reproductive process. For example, a protocol can be developed to ascertain fertility, serum and tissue hormone levels, testis histopathology, testis daily sperm production, epididymal sperm reserves, sperm morphology, sperm motility, all in the same animal.

SUMMARY

The proposals of shorter-term testing protocols impact the manner in which histopathology is conducted. Histopathology utilizing upgraded tissue preparation techniques conducted by a highly trained histopathologist will be necessary to detect many adverse effects. Development of rapid quantitative techniques, preferably computer automated techniques, will allow the toxicologist to

statistically evaluate counts of cells to either show a "no effect" or to show an "effect" and, in the latter case, to provide objective data showing the severity of the effect.

Histopathological evaluation of the testis is not conducted "in a vacuum" in a toxicology testing setting. There is considerable other information about the agent that can guide and influence the testing protocol to increase the sensitivity of histopathological evaluation.

REFERENCES

Amann, R.P. 1970. Sperm production rates. In: Johnson A.D., Gomes W.R., and Van Demark N.L., eds., The Testis, Academic Press, N.Y., pp. 433–482.

Anderson, R.J., Berryman, S.H., Phillips, J.F., Feathergill, K.A., Zaneveld, L.J., and Russell, L.D. 1989. Biochemical and structural evidence for ethanol-induced impairment of testicular development: apparent lack of Leydig cell involvement. *Toxicol. Appl. Pharmacol.* 100:62–85.

Berndtson, W.E. 1989. Sampling intensities and replication requirements for detection of treatment effects on testicular function in bulls and stallions: a statistical assessment. *J. Anim. Sci.* 67:213–225.

Berndtson, W.E. 1977. Methods for quantifying mammalian spermatogenesis: A review. *J. Anim. Sci.* 44:818–33.

Blazek, W.F., Rushbrook, C.J., Ernst, T.L., Stewart, B.E., Spak, D., Di Biasio-Erwin, D., and Black, V. 1985. Relationship between breeding performance and testicular/epididymal functioning in male Sprage-Dawley rats exposed to nitrobenzene (NB). *Toxicologist* 5: 121 (abstract).

Blazak, W.F., Treinen, K.A., and Juniewicz, P.E. 1993. Application of testicular sperm head counts in the assessment of male reproductive toxicity. In: Chapin R.E., Heindel J.J., eds. Male Reproductive Toxicology. Methods in Toxicology; Vol. 3A, Academic Press, San Diego: pp. 86–94.

Chapin, R.E. 1988. Morphological evaluation of seminiferous epithelium of the testis. In: Lamb J.C. IV, Foster, P.M., ed., *Physiology and Toxicology of Male Reproduction.* Academic Press California, pp. 155–173.

Chapin, R.E., Ross, M.D., and Lamb, J.C. 1984 Immersion fixation methods for glycol methacrylate-embedded testes. *Toxicol. Pathol.* 12:221–227.

Chapin, R.E., Russell, S.R., Cocanougher, T.A., and Gulati, D.K. 1992. Testicular spermatid head count (SHC) compared to other indices of reproductive toxicity: An interim report. *Toxicologist* 12:433 (abstract).

Evenson, D.P., Higgens, P.H., Greneberg, D., and Ballachey, B. 1985. Flow cytometric of mouse spermatogenic function following exposure to ethylnitrosourea. *Cytometry* 7:45–53.

Evenson, D.P. and Melamed, M.R. 1983. Rapid analysis of normal and abnormal cell types in human semen and testis biopsies by flow cytometry. *J. Histochem. Cytochem.* 31:248–253.

Harris, M.W., Chapin, R.E., Lockhart, A.C., and Jokinen, M.P. 1992. Assessment of a short-term reproductive and developmental toxicity screen. *Fund. Appl. Toxicol.* 19: 186–196.

Hess, R.A. and Chen, P. 1992. Computer tracking of germ cells in the cycle of the seminiferous epithelium and prediction of changes in cycle duration in animals commonly used in reproductive biology and toxicology. *J. Androl.*13:185–190.

Hess, R.A. and Moore, B.J. 1993. Histological Methods for Evaluation of the Testis. In: Chapin, R.E. and Heidel, J.J., eds. Methods in Toxicology. Vol 3A. Academic Press, San Diego. pp. 52–85.

Ku, W.W., Chapin, R.E., Wine, R.N., and Gladen, B.C. 1993. Testicular toxicity of boric acid (BA): Relationship of dose to lesion development and recovery in the F344 rat. *Reprod. Tox.* 7:305–319.

Robb, G.W., Amann, R.P., and Killian, G.J. 1978. Daily sperm production and epididymal sperm reserves of pubertal and adult rats. *J. Reprod. Fert.* 54:103–107.

Russell, L.D. 1983. Normal testicular structure and methods for its evaluation under experimental and disrupted conditions. In: Clarkson, T.W., Nordberg, G.F., and Sager, P.R., eds, Reproductive and Developmental Toxicity of Metals. Plenum Press, NY. pp. 227–252.

Russell, L.D. 1993. Role in Spermiation. In: Russell L.D., Griswold M.D., eds. The Sertoli Cell. Cache River Press, Clearwater, FL pp 269–303.

Russell, L.D., Ettlin, R.A., Sinha Hikim, A.P., and Clegg, E.D. 1990. Histological and Histopathological Evaluation of the Testis. Cache River Press, Clearwater, FL. p. 280.

Wing, T-Y. and Christensen, A.K. 1982. Morphometric studies on rat seminiferous tubules. *Am. J. Anat.* 165:13–25.

Endnote

[a] This report will not specify nor cite any particular protocol in use or being considered for use since the concepts presented should be applicable to any protocol, current or future.

REFINEMENTS IN THE EXPOSURE ASSESSMENT PROCESS

Michael L. Gargas[1], Paul K. Scott[1], Brent D. Kerger[2], Brent L. Finley[1], and Richard H. Reitz[1]

[1]ChemRisk, Division of McLaren/Hart
The Courtland East Building, 29225 Chagrin Blvd., Suite 300, Cleveland, OH 44122

[2]ChemRisk® Division of McLaren/Hart, Irvine, CA

REFINEMENTS IN THE EXPOSURE ASSESSMENT PROCESS

Some of the most significant advances and refinements in human health risk assessments for environmental and occupational chemicals have come from investigations in the area of exposure assessment. This presentation summary describes three previously published examples of continuing advances in this process.

In the first example a physiologically based pharmacokinetic (PBPK) model was used to predict acceptable levels of human exposure to chloroform in tap water (Kerger et al., 1994). The modeling supported an alternative Maximum Contaminant Level (MCL) that considered oral, dermal, and inhalation uptake of chloroform during household exposures. Three exposure scenarios were evaluated. Scenario I was quite conservative (i.e., 24 hours/day spent indoors, 2L of water ingestion/day consumed at three time periods, continuous light exercise); Scenario II was less conservative (i.e., 14 hours/day spent indoors, 1L of water ingested throughout the day, sedentary life-style); and Scenario III was more realistic (i.e., 14 hours/day spent indoors, sedentary except for 30 minutes of heavy exercise, 0.33L of water per meal and 1L immediately after heavy exercise). The results of the PBPK modeling for the three scenarios indicate that upper bound cancer risks from tapwater-related exposures to chloroform may be considered negligible (e.g., ranging from 0.3 to 3 potential extra cancers per million) at tap water concentrations up to 2000 mg/L. Chloroform exposures at the current USEPA MCL for trihalomethanes (100 mg/L, assuming all chloroform) are well below a de minimis risk level based on the PBPK model for liver cytotoxicity and tumorigenicity.

The second example demonstrated the benefits of using a probabilistic exposure assessment (Finley, et al., 1994). In this study, the risks associated with household exposure to either bromoform, chloroform, tetrachloroethylene, or vinyl chloride, in tapwater at concentrations equal to current MCLs

were evaluated using both point estimates for the Reasonable Maximal Exposure (RME) and Population Density Functions (PDFs) in a probabilistic assessment using Monte Carlo simulations. The exposure pathways evaluated were direct tapwater ingestion, dermal contact while showering, indoor inhalation of vapor, and ingestion of garden vegetables irrigated with contaminated water. It was shown that the RME approach overestimated risk for bromoform, chloroform, tetrachloroethylene, and vinyl chloride by factors of 1.5, 8.2, 1.4, and 1.9, respectively, over the probabilistic approach when comparing the 95th percentile person. This study concluded that MCL concentrations of these chemicals are unlikely to pose a significant cancer risk ($> 10^{-4} - 10^{-6}$) to a resident who uses tapwater for drinking, showering, and garden irrigations.

The last example described a reevaluation of benzene exposure among pliofilm workers (Paustenbach et al., 1992). This cohort has been extensively investigated as part of several earlier studies. Eight exposure factors were considered in the re-analysis: (1) Short term, high-level exposures; (2) Background concentrations; (3) Dermal contact; (4) Updated morbidity data; (5) Installation and effectiveness of IH engineering controls; (6) Extra-long workweeks during the war years; (7) Shortcomings in the air sampling and monitoring devices; and (8) Effectiveness of respirators and gloves. New historical information on short-term high-level exposures as well as consideration of dermal absorption (including evaluations using Monte Carlo analysis) demonstrated that past evaluations underestimated the total absorbed dose to some workers by 25–50%. These findings will have implications for estimating future cancer potencies for benzene in people.

Each of the refinements in exposure assessment described above help to decrease the uncertainty inherent in risk assessments. Tools such as PBPK modeling, Monte Carlo simulation, and retrospective analyses allow evaluations that need not rely so heavily on conservative, default assumptions, and allow a more realistic assessment. Future risk assessments should rely more and more on such tools.

REFERENCES

Finley, B.L., Scott P., and Paustenbach, D.J. (1993). Evaluating the adequacy of Maximum Contaminant Levels as health-protective, cleanup goals: an analysis based on Monte Carlo techniques. *Regulatory Toxicology and Pharmacology*, 18:438–455.

Kerger, B.D., Walker, L.B., Gargas, M.L., Paustenbach, D.J., and Reitz, R.H. (1994). Perspective on risks of chloroform in tapwater: Implications of physiologically based pharmacokinetic modeling on regulatory toxicity criteria. *Toxicologist,* 14:153.

Paustenbach, D.J., Price, P.S., Ollison, W., Blank, C., Jernigan, J.D., Bass, R.D., and Peterson, H.D. (1992). Reevaluation of benzene exposure for the pliofilm (rubberworker) cohort (1936–1976), 0, 36, 177–231.

SESSION V

CHARACTERIZATION OF RISK IN A TEMPORAL CONTEXT

ESTIMATING HEALTH RISK IN OCCUPATIONALLY EXPOSED NAVY PERSONNEL

John F. Risher, Warren W. Jederberg, Robert L. Carpenter

Navy Medical Research Institute
Toxicology Detachment, Wright Patterson Air Force Base, Fairborn, Ohio

ABSTRACT

Occupational exposures of Naval personnel to toxic chemicals and other substances occur on a recurring basis by virtue of their jobs and the mission of the Navy. Such exposures are similar to, yet in many respects distinct from, either occupational exposures of civilian workers or environmental exposures of the general population. Primary dissimilarities may include the duration of exposure, the intermittent nature of exposure, and the age, health, and physical condition of the exposed population. Although the duration of individual exposures may be similar, the total exposure of a worker over a period of months or years is likely to be different. Any exposures at a given duty station are often intermittent, with prolonged (days or months) exposure-free periods in between. The Navy, therefore, is faced with developing an occupational risk assessment paradigm specific to, and appropriate for, the types of occupational exposures experienced by Navy personnel in the performance of shipboard or other duties. Acceptable risk must take into account not only the protection of worker health, but also the ability of the exposed individual to fulfill the Navy mission at hand. This may mean, in some situations, that the toxicity endpoints used as a basis of chronic risk determination for the general population may not be the most appropriate for Navy personnel. This paper discusses the differences between Navy and non-military exposure scenarios, contrasts the merits of some of the existing methodological approaches to estimating health risk in a Navy setting, and discusses the need for a centralized data base for medical surveillance of occupationally exposed Navy personnel.

INTRODUCTION

Occupational exposures of Naval personnel to toxic chemicals and other substances occur by virtue of both an individual's duties and the mission of the Navy. Such exposures are similar to, yet distinct from, either occupational exposures of civilian workers or environmental exposures of the general population. While environmental exposures (e.g., to substances at or emanating from a hazardous

waste disposal site) are assumed to be for up to 30 years in duration for non-carcinogenic substances and 70 years for carcinogens, and civilian occupational exposures are typically for 8 or 10 hour days and 40-hour work weeks for up to 50 years (NRC, 1993), exposures of Navy personnel on the job are often more limited and of an intermittent nature.

The Navy, therefore, is faced with developing an occupational risk assessment paradigm specific to, and appropriate for, the types of occupational exposures experienced by Navy personnel in the performance of shipboard or other (rate/watch station-specific) duties. It is not necessary, or perhaps even appropriate, that this paradigm be different than those peer-reviewed processes used by the U.S. Environmental Protection Agency (EPA), National Institute for Occupational Safety and Health (NIOSH), Agency for Toxic Substances and Disease Registry (ATSDR), the Occupational Safety and Health Administration (OSHA), and the World Health Organization (WHO). The needs of the Navy might, in fact, be satisfactorily met by utilization of existing risk and exposure assessment protocols where appropriate and applicable, and modification or improvement of them to be representative of the exposure scenarios confronting Naval personnel in non-combat occupational scenarios.

To develop such a protocol, it is first necessary to evaluate the similarities and differences between the Navy occupational, civilian work force, and involuntary environmental exposures. A certain amount of this can be done using knowledge of both common and exceptional/atypical workplace practices and exposed populations, while some aspects of such a determination must involve assumptions concerning hypothetical exposure scenarios and various parameters relevant to the assessment of health risk (e.g., 70 kg reference adult, 20 m^3 of air inhaled per day). A knowledge of the various factors affecting the assessment of risk among different exposed populations must be considered prior to any quantitative risk determination based upon the same or similar data. Similarities and differences in duration of exposure, frequency of exposures, intensity (dosage) of exposure, and age, sex, and general health of the exposed population are illustrated in Table V-1.

Navy personnel are typically in better health and physical condition than the corresponding civilian work force. The Navy has a requirement for twice-a-year physical readiness (fitness) testing and stringent height-weight-body fat requirements that preclude a service member who cannot meet these standards for longer than a short-term, temporary basis from remaining in the Naval service.

Another issue is that of average age of a Navy man or woman compared with his or her civilian work force counterpart. The mean age of the Navy work force is 27 years (Conway et al., 1989), compared with 35 years for the overall civilian work force (U.S. Census Bureau, 1993). Further, a

typical career Navy enlisted person would be expected to be on active duty between the approximate ages of 18 and 38 to attain the necessary 20-year longevity for retirement, in sharp contrast to the Social Security retirement age of 65 for private sector employees not under corporate or union pension plans.

Table V-1. Comparison of Exposure Parameters Among Different Groups.

Parameter	Navy	Civilian Occupational	Environmental Involuntary
Nature of Exposure	Intermittent to Subchronic	Intermittent to Subchronic	Assumed to Be Continuous
Length of Exposure	May (rarely) Be > 8 Hours/Day, but Typically Discontinuous/ Intermittent (due watch rotation, transfers, etc.)	8–10 Hours/Day, 40 Hour Work Week	Daily Oral/ Inhalation (w/assumed quantity ingested/ inhaled)
Frequency of Exposure	Up to 7 Days/Week, but Typically 5 or Less Days per Week	Daily (5 days /week); up to 40+ Years	Daily/Continuous
Route of Exposure	Primarily Inhalation and Dermal	Primarily Inhalation and Dermal	Primarily Oral; Inhalation Secondarily; Dermal not Quantitatively Assessed, but Assumed to be Minimal
Age of Exposed Population	17–45	18–60	Birth–70+
Gender of Exposed Population	Both (primarily male)	Both (primarily male)	Both
Physical/ Medical Condition of Exposed Population	Excellent*	Variable, but Generally Good	Various (range from excellent to feeble or high risk)

* Selected against adverse health and excessive % body fat.

The exposure scenario in a Navy occupational environment is typically different than that in an industrial or other non-Navy setting. Although the duration of individual exposures may be similar, the overall time in which a worker is exposed is likely to be different. Navy exposures are typically either limited to a single tour of duty (1½ to 3 years), during which exposure occurs, or interrupted/separated

by periods of years with no exposure. Further, any exposures at a given duty station are often intermittent, with prolonged (days or months) exposure-free periods in between.

Thus, exposures to toxic substances in the Navy are to a typically younger, healthier, more physically fit work force for a significantly shorter period of time. Consideration of these factors and of the differences in the nature of exposure(s) taken collectively suggest that a toxicity-based health risk assessment paradigm for Navy personnel might be somewhat different than that used in civilian occupational settings or for involuntary environmental exposures of the general population (all age groups included).

COMPARISON OF EXISTING METHODS

The issue of occupational health risk is not new, and has previously been addressed by other federal agencies. The National Institute for Occupational Safety and Health (NIOSH) is charged with conducting health studies to determine and assess health risks associated with occupational exposures to potentially toxic substances in the workplace. NIOSH not only conducts laboratory experiments using animals as surrogates for mammalian toxicity, but conducts occupational exposure assessments and develops human health criteria documents, as well. The Occupational Safety and Health Administration (OSHA) is charged with setting occupational standards in the workplace, and sets health-based limits (e.g., permissible exposure limits, or PELs) for workplace exposures for healthy adults. Also in the occupational area, the American Council of Governmental Industrial Hygienists (ACGIH) publishes annually a list of time-weighted-average threshold limit values (TLV-TWAs) for use in protecting workers from excessive exposure to toxic substances. [These TLV-TWA values are intended to protect healthy adult workers, and are based upon 8 hour/day, 40 hour/week exposure scenarios. They do not necessarily reflect the most recent scientific data, however, and are not used by the U.S. Environmental Protection Agency (EPA) and Agency for Toxic Substances and Disease Registry (ATSDR) as a basis of determining health protective levels of exposure for the general population or in all occupational settings.] The EPA (a federal regulatory agency) and ATSDR (under the U.S. Public Health Service) establish health-based guidance values that apply to the general population, including all age groups and individuals of particular sensitivity/susceptibility to the toxic effects of hazardous substances.

While these federal agencies have developed presumed protective exposure levels for many chemicals, minerals, and other substances for the different environmental media (particularly air and water), the exposure scenarios employed and the nature of the populations considered are not

necessarily fully relevant to occupational exposures of Navy personnel. Further, the Navy has some chemicals [e.g., torpedo (OTTO) fuel and DBMP] which are unique to the Navy since they are manufactured for, and used exclusively by, the Navy.

The Committee on Toxicology of the National Research Council has developed Criteria and Methods for Preparing Emergency Exposure Guidance Levels (EEGLs), Short-Term Public Emergency Guidance Levels (SPEGLs), and Continuous Exposure Guidance Levels (CEGLs) for some chemicals of concern to the Department of Defense (DoD) (NRC, 1986). While the emergency exposure levels developed by NRC are relevant to some Navy scenarios, they do not specifically or fully cover the exposure of Navy personnel during normal occupational conditions.

Emergency Exposure Guidance Levels (EEGLs) represent peak concentrations of substances in air (as a gas, vapor, or aerosol) that may be judged by DOD to be acceptable for the performance of specific tasks during rare emergency conditions lasting for periods of 1-24 hours (e.g., during a fire, chemical spill, line break, etc.). Exposure at an EEGL might produce reversible effects that, although not necessarily hygienic or safe, do not impair judgment or interfere with proper responses to the emergency.

EEGLs are based upon the acute toxicity of the substance in question. They are intended to cover all effects, immediate and delayed. They are based upon the assumption that the exposure will be followed by complete recovery. In the case of carcinogens, upper limit lifetime cancer risk will be no greater than 1 in 10,000 (10^{-4} risk), recognizing that if exposures can be limited to a few times or less, cancer is less likely to develop.

An EEGL is acceptable only in an emergency, when some risks or discomfort must be endured to prevent greater risks (e.g., fire or explosion). Exposure at the EEGL might produce effects such as increased respiratory rate from increased CO_2, headache or mild CNS symptoms from CO, or respiratory tract or eye irritation from ammonia, phosgene, or SO_2. The intent of the EEGL is to prevent irreversible damage; and while some reduction in performance is considered permissible under these conditions, the effects of exposure should not prevent proper responses to the emergency (e.g., shutting off a valve, closing a hatch, using a fire extinguisher, etc.). While in normal working situations, upper respiratory tract or eye irritation causing discomfort would not be considered acceptable, the same symptoms would be considered acceptable under emergency conditions, if they did not cause irreversible harm or seriously affect judgment or performance.

In the development of EEGLs, safety (uncertainty) factors are typically employed. However, the safety factor of 10 for extrapolation from laboratory animals to humans is the most commonly, and may be the only, one used.

Continuous Exposure Guidance Levels (CEGLs) are ceiling concentrations intended to avoid adverse health effects, immediate or delayed, and to avoid degradation in the performance of military personnel after exposure for up to 90 days. Parameters considered in deriving CEGLs that are not considered in EEGL derivation include accumulation, detoxification, and excretion. These guidance levels are not intended for use as community or work standards, but are intended to estimate the effects of exposures during operations lasting up to 90 days in an environment like that in a submarine.

CEGLs are generally set at 0.01 to 0.1 times the 24-hour EEGL (e.g., with safety factors of 10 to 100). For example, if there is evidence of substantial detoxification, a safety factor of 10 might be sufficient. However, if there is no or limited (i.e., slow) evidence of detoxification, a safety factor of 100 might be more appropriate. Even greater safety factors may be considered necessary in the case of substances that accumulate in the body (e.g., metals and halogenated biphenyls).

Short-term Public Emergency Guidance Levels (SPEGLs) are defined as suitable concentrations for unpredicted, single, short-term emergency exposure of the general public. Unlike the EEGL and CEGL, the SPEGL takes into account sensitive populations (e.g., children, aged, people with serious medical conditions or debilitating diseases). SPEGLs are generally set at 0.1 to 0.5 times the EEGL. (A safety factor of 2 is generally considered by the NRC Committee on Toxicology to be appropriate to protect more sensitive groups, such as children and the elderly, and a factor of 10 is considered appropriate to protect fetuses and newborns.)

While the EPA, ATSDR, NIOSH, and OSHA health guidance values developed to protect the civilian population would typically be expected to also be protective of Navy personnel under appropriate exposure scenarios, they might be unnecessarily stringent for application to other Navy scenarios. The reasons for this are threefold. First, the typical Navy sailor represents a healthier individual than a typical non-military counterpart. This is supported by the Navy's pre-enlistment medical screening of prospective enlistees to eliminate those with serious or chronic health problems, screening against unfit or overweight personnel through mandated semi-annual physical readiness tests (PRTs), and having rigid gender-specific height-weight and body-fat standards which all individuals, regardless of age, rate, or rank are required to meet. Failure to attain these standards will result in involuntary separation from the Naval Service. The effectiveness of this program is attested by the

lower body fat percentage compared with the civilian working population. Body fat measurements reported by Conway et al. (1989) for Navy males averaged approximately 16%, compared with 19% for non-military males. Average body fat measured in Navy females in that study was just under 24%, but no comparable figures were available for non-military females. Body fat was, however, found to increase with age in both sexes. The disparity in body fat between the Navy and civilian work forces might be even greater if specific occupational subgroups, such as Navy divers or SEALs, are considered separately. [For example, the average body fat composition of Navy males recorded prior to physical readiness testing was reported to be 16% (Conway et al., 1989); this may be contrasted with 10.4% for Navy BUDS (divers in training to be SEALs) reported by Flemming et al. (1993).] The Navy also has a continuing drug screening program, HIV screening program, and a requirement for periodic physical examinations to ensure the health and fitness of all personnel.

Secondly, the average age of the typical sailor is less than the age of the corresponding industrial worker. The mean age of the Navy work force is 27 years (for both males and females; Conway et al., 1989), compared with 35 years for the overall civilian work force (U.S. Census Bureau, 1993). A person who comes into the Navy at age 18 and does not progressively or substantially advance in rank, but who nonetheless makes it a career will probably (based upon historical observation) retire at approximately 38 to 40 years of age (i.e., after 20 years of service). A person on active duty beyond age 38 would typically be in a senior, supervisory position, and be unlikely, or at least far less likely, to be continuously exposed to any toxic chemical in a peacetime environment. By contrast, a private sector/civilian worker who begins a line of work at age 18 and does not advance or aspire to a supervisory level may remain (with assumed pay raises for experience, etc.) until corporate/union pension or Social Security (65) retirement age, incurring chronic exposure to a chemical agent or agents during that period.

A third factor that sets Navy occupational exposures apart from private sector exposures is the overall length of exposure. The exposure scenario experienced by most Navy men and women is of shorter duration [up to 3 years total in most cases vs. up to 50 years total (NRC, 1993)] than their civilian counterparts. Unlike the industrial worker who may have the same job for 30 or 40 years, the exposure of a Navy worker is of a comparatively briefer (1 to 3 years) timespan, with any subsequent period of similar exposures likely to be interrupted by periods of duty in another area (i.e, transferred to another ship or shore station where daily duties might be substantially or totally different from an exposure perspective). At a given duty station, exposures to potentially toxic substances are often

intermittent, with exposure periods punctuated with prolonged exposure-free periods of days or months. In addition, the career path within a Navy rate (job specialty/classification) includes both periods of changes in individual duties (including administrative assignments) and changes in command/duty station. It is therefore highly unlikely that a Navy member in most, but not all, ratings and work scenarios would confront the truly chronic type of occupational exposures seen in the private sector[a]. It is acknowledged that there are certain ratings, such as hull technicians (HTs), that may inherently involve exposures of a more chronic nature to some chemicals/substances; however, there are already existing workplace exposure standards and/or methodological approaches to health risk determination that may be applied to such exposures.

In the case of health-based guidance values established by EPA and ATSDR for inhalation exposures, such values are typically derived from either studies of occupational exposures or clinical/controlled studies of human subjects or laboratory animals with limited daily exposures and, particularly in the case of human studies, sub-chronic or shorter duration. Since EPA and ATSDR are charged with the protection of the health of the overall population under continuous exposure conditions, adjustments are typically made to the dose/concentration levels used in published studies (human and animal) to extrapolate to a 24-hour a day, seven day a week exposure scenario (MRL; U.S. EPA, 1990). Such an adjustment may not be necessary in the case of Navy occupational exposures.

The exposure of Navy personnel is further differentiated from occupational exposures of the civilian work force by the career path of a worker. Whereas a pesticide formulator, solvent mixer/manufacturer, dry cleaner, steel worker, or other worker may perform the same or similar tasks on a daily basis for decades, a Navy sailor is expected to advance in grade (and technical expertise) to a petty officer, or supervisory level, at which hands-on activity/behavior resulting in exposure to toxic substances is typically reduced with each successive promotion. Further, as a sailor advances in pay grade during his/her career, the career path is punctuated with periods of administrative-type duty where chemical exposure of any kind is either unlikely or minimal. This may be contrasted with a typical private sector industrial worker, who is paid to perform a specific task or tasks, and for whom advancement to a supervisory or less technical managerial or administrative position is usually not a corporate priority. In the Navy, if a person shows little or no promise for advancement (promotion), that person may eventually be separated from the service to make room for individuals more compatible with Navy needs and lifestyle.

ROUTES AND DURATION OF EXPOSURE

In terms of actual Navy exposure scenarios, four distinct scenarios are deemed to be relevant to occupational exposures: (1) chronic, defined as exposures equal to or greater than one year in duration (i.e., daily or multiple times per week exposures for extended, uninterrupted periods of time); (2) subchronic/intermediate, considered to be greater than acute, but less than chronic exposure; (3) acute, defined as consecutive daily exposures equal to or less than 14 days total duration; and (4) intermittent, considered to be a series of acute exposures, separated by a period of 14 days or greater (may be more or less than 14 days, based upon toxicokinetic parameters of the particular chemical or substance in question).

The routes of exposure in Navy occupational settings, as with non-military workplace exposures, are primarily inhalation and dermal, with inhalation being the primary method of exposure in most instances. Depending on the substance in question, however, oral could be a route of exposure under conditions of inadequate personal hygiene. For example, in the case of exposure to asbestos or lead-containing dusts in the workplace, an individual who does not wash their hands properly before eating and/or does not change their clothing when leaving work runs the risk of significant oral exposure, as well as the possibility of transporting/tracking the contamination to other areas, including the home or other residence, which might result in exposures to other individuals as well. Nonetheless, most exposures involving Navy personnel will be either inhalation and/or dermal in nature.

ENDPOINTS

Another issue to be addressed is whether toxicity endpoints selected for Navy occupational exposures should be the same as those used by EPA and ATSDR for the general population. The differences in the general health of the Navy and overall U.S. populations, the narrower range of potentially exposed age groups (i.e., 18–40+ for Navy vs. infants to the aged in the general population), and the shorter duration of overall exposure periods (chronic effects vs. shorter-term effects) suggest that endpoints for some Navy exposure scenarios might be different from those used for a general population exposure scenario.

For U.S. regulatory agencies, such as EPA, decisions regarding environmental standards are frequently based upon cancer endpoints. The scientific bases for these standards are typically chronic, lifetime cancer bioassays in one or more mammalian species (e.g., may be human or other primates, but more typically rodents). Studies involving laboratory animals typically use very high dosages [e.g., 1.0 and 0.5 times the maximum tolerated dose (MTD)] to maximize the probability that tumors will

result from the experimental exposures in a relatively small animal population. The multi-stage model used by the EPA provides for the extrapolation of the results of such a study (-ies) to a low-dose, assumed human lifetime exposure of 70 years. These calculations are typically used for estimation of an exposure dosage/air concentration that would result in an upper limit risk of one additional cancer in a population of a given size [usually one million (or 10^{-6} upper limit risk level), although 10^{-5} or 10^{-4} levels may be used in the case of environmental clean-ups at Superfund sites (EPA, 1989) (a risk management decision)].

While such cancer-based regulatory standards would be expected to be protective of lifetime exposure by Navy personnel as well as the general civilian population, Navy personnel are not exposed for a 70-year lifetime. Rather, typical exposures to solvents or other substances considered to be carcinogens are most likely to be intermittent in nature and often of 3 years or less duration for total workplace exposures for most Navy ratings. Prorating/adjusting the cancer dosage for lesser durations of time has been used by some federal and state/local environmental/health agencies in an attempt to use the assumptions of the multi-stage model for shorter-term exposure scenarios; however, this creates additional uncertainty about the ultimate calculated number. Given that the specific focus of the 1994 Conference at which this paper was presented was stated to be non-cancer endpoints, no further discussion of this area will be presented here.

Much of the published literature on mammalian health effects of exposures to toxic substances identifies toxicity endpoints of questionable relevance to Navy personnel. For example, the only non-cancer effect reported in a chronic (e.g., 2-year) rat feeding or inhalation bioassay might be a slight reduction in body weight gain or a change in organ weights without accompanying biochemical perturbations or histopathological changes in the same organ(s). While an argument can be made for the undesirability of such weight changes, the biological significance of those effects is less clear. The impact of such subtle effects as body and organ weight changes by themselves on the overall health or performance of intermittently exposed, young, physically fit, healthy workers may not be considered sufficient to serve as a basis for establishment of health-protective exposure guidance limits without at least supporting pharmacokinetic or structure-activity information and through examination of the relevance of the exposure scenario vis a vis the projected Navy occupational exposure scenario. For example, the relevance of effects observed in laboratory animals during daily chronic exposures may not always be extrapolatable to intermittent human exposures, depending upon the toxicokinetic properties of the substance in question. This is illustrated by the observation of hepatotoxic effects in

mice exposed to JP-4 jet fuel vapors at a concentration of 500 mg/m^3 for 24 hours per day, 7 days a week for 90 days (U.S. Air Force, 1984), but not in mice exposed for only 6 hours per day, 5 days a week for 8 months at an air concentration of 5,000 mg/m^3 (U.S. Air Force, 1974) or for 6 hours per day, 5 days a week for 1 year at 1,000 mg/m^3 (Bruner et al, 1991).

On the other hand, since many shipboard duties require complete alertness, ability to discriminate between variables or objects (e.g., radar, electronics sensory measures, or sonar operator, lookout, or air controller) and full presence of all physical capacities (e.g., balance, hearing, etc.), studies which identify obtunding of the senses or other neurological/neurobehavioral effects would be particularly relevant. There are a number of substances (e.g., the solvents toluene and xylene) for which human data in this area are already available, but clinical evaluation of exposed Naval personnel to identify the presence of toxic effects, and correlate the observed effects with relevant exposure levels and durations, could significantly enhance the toxicity data bases for those and other substances and facilitate the determination of levels of toxicants likely to produce adverse health effects in humans. This is particularly true of substances known to produce discernable neurological and other more subtle effects in laboratory animals, but for which human data are lacking.

The duration of exposure of toxicological studies should be comparable to those of the exposed Navy population. Typically, subchronic studies ranging in duration from 90 days to 1 year would be the most desirable. Such studies should examine a number of parameters, including changes in gross appearance and behavior, periodic examination of clinical and biochemical indices, and gross and microscopic histopathological examination of all major organs. Reproductive and developmental effects would also be highly relevant, although such studies are typically conducted separately from the type of subchronic assays needed. A type of study that is rarely available is one (esp. inhalation) in which the test subjects were exposed on an intermittent basis (e.g., 6 hours/day, 3 days per week) with examination of pharmacokinetic/toxicokinetic parameters as well as classical toxicological endpoints. If such a study examining the same endpoints identified in traditional longer-term studies, but using an intermittent exposure scenario, were to be conducted and compared with the results of the traditional 90-day or chronic study using the same dose or air concentrations, it could provide potentially invaluable scientific data and reduce the uncertainty in any estimates of adverse health potential associated with intermittent exposures to the same chemical/substance. Table V-2 provides the "ideal" data base for use in assessment of the health risk associated with Navy occupational exposures.

Table V-2. Ideal Data Base for Assessing Risk in Navy Occupational Exposures.

Human	or	Animal
1. Clinical Studies (controlled) • acute • subchronic/ intermediate • variety of endpoints including neurobehavioral indices (e.g., nerve conduction velocity, performance tests, etc. appropriate for chemical) 2. Epidemiological Study • occupational • other prospective or retrospective study		1. One 90-day Study* Using Appropriate Species (usually rodent), looking at: • body weight changes • clinical/serum parameters • biochemical indices • gross observations • gross and microscopic histopathology 2. Second 90-day (or longer subchronic or chronic) Study in a Second Species 3. Intermittent Exposure Study (e.g., 3 days per week, 6 hours per day) in Animal Species with Appropriate Sensitivity for the Primary Effect of Concern (as determined from previous ani and/or human studies) 4. Rodent Reproductive Toxicity Study (preferably multi generation) 5. Developmental Study (rabbits or rodents) 6. Neurobehavioral Toxicity Study

* A single 90-day subchronic study examining, as a minimum, these parameters represents a minimum data base for Navy occupational risk assessment.

BIOMARKERS

Biomarkers of exposure, effect (harbingers of the onset of toxicity), and susceptibility can play a key role in facilitating the identity of the "crucial uncertainty," i.e., how do we know that someone is really being, or has been, significantly exposed? A good biologic marker not only should be easily measured and directly related to the toxicant, but it should also be related to it in a known way (i.e., in a way that fits the marker with the mechanism of toxic effect). The dose/exposure conditions at which the presence of the biomarker is first observed should be contrasted with the dose and level of biomarker present at the onset of toxicity.

Biomarkers can serve to identify individuals who are particularly at risk, due to particular sensitivity to a potentially toxic substance. If not already available in the medical literature or published toxicity studies, the Navy can play a significant role in identifying markers for known levels of

exposure with the aim of extrapolating for predictive purposes to people with unknown levels of exposure. In the case of known biomarkers, the gathering of clinical data from Navy personnel known to have been, or suspected of having been, exposed to a toxic substance in the workplace can assist in tying a concentration of biomarker with the onset of manifestation of toxicity in human subjects. Ideally, they should be chemical-specific, particularly in the case of multiple chemical exposures. Biomarkers can range from non-toxic metabolites of a parent compound in the urine to actual identification of a known toxic metabolite in breast milk or blood.

USE OF REGISTRIES

ATSDR, a component of the U.S. Public Health Service, is tasked by Congress with establishing and maintaining disease/toxicity registries to track health effects of populations exposed to hazardous substances at Superfund and other hazardous waste disposal sites meeting certain criteria. These registries are not only useful for tracking the health of exposed populations at one specific site, but provide information concerning health effects that may be used for predictive purposes, as well. Although the Navy has acknowledged the need for medical surveillance of its occupationally exposed personnel (DoD, 1984; OPNAV, 1983) and has an organizational framework suitable for establishing and maintaining permanent registers for occupational exposures of its personnel, the merit of a centralized reporting and tracking system for occupational exposures has here-to-fore not been fully recognized. The use of registries or a similar central reporting and tracking function/system under the control of a centralized reporting authority (e.g., the Navy Bureau of Medicine and Surgery, or BuMed) would serve as an invaluable source of information to the Navy medical community, and significantly enhance the scientific data base of human occupational exposures and risk assessments, as well. Comparison of such Navy-wide registries with ATSDR and any other similar tracking registries could serve not only as a cross-check, but would also substantially facilitate the understanding of the effects of toxic substances on exposed humans.

NAVY MEDICAL SURVEILLANCE PROGRAMS

Medical surveillance programs for monitoring the health of Navy personnel potentially exposed to toxic substances by virtue of either their military rating/job or of known hazards in the workplace have been in existence for some time. Existing guidance (DoD, 1984; OPNAV, 1983) states that the selection of personnel for medical surveillance examination should be based primarily upon the results of industrial hygiene (IH) surveys that quantify exposure in the workplace (i.e., "hazard-based

surveillance"). In many instances, however, workers are also placed into surveillance programs based upon their job or billet (e.g., waste workers of firemen) (NMC, 1988).

Workers whose jobs are associated with exposure to hazards or stressors greater than an established action level for an identified substance or stressor for greater than 30 days per year or 10 days per quarter are placed in medical surveillance programs. When no action level or legal standard for medical exams for specific agents exist, then one half or more of a recommended exposure limit (e.g., ACGIH TLV-TWA) may be used as the action level (DoD, 1984).

The decision to place a Navy worker in a medical surveillance program should be made by the best qualified medical department representative available. This may be a trained occupational medicine physician, a trained occupational nurse, or an industrial hygienist; however, all of these may in fact be involved in the decision.

Navy IH personnel sampling strategies and data obtained from sampling are well-suited for workers involved in static, on-going processes (e.g., a plating operation that occurs 40 hours/week, 52 weeks/year). However, many Navy workers perform operations which may place them in exposure situations only several times per week, month, or year, making it difficult to collect IH data representative of the operation. Nonetheless, the medical surveillance program is a source of empirical, quality data which, when coupled with known exposure information, provide both a source of protection for the worker and a valuable source of human exposure-effect data relevant and applicable to exposure of the civilian work force and general public, as well.

The Navy Environmental Health Center (NEHC, 1991) lists over 100 chemical and physical stressors to which exposure may warrant participation of a worker in the Navy's medical surveillance program. These stressors are listed in Table V-3. Although the surveillance program represents a wealth of exposure, biomarker, and health effect information, there is at yet no coordinated framework to provide a centralized permanent record for an individual (aside from his/her service medical record, a private record which accompanies a service member from duty station to duty station), to provide a composite picture of exposure to multiple stressors throughout a career, or to provide a collective record of all individuals exposed to a particular stressor. Establishment of a permanent, centralized record of exposures to Navy personnel which would couple information on the length of individual exposures and the duration of overall exposure with observed effects would be of considerable value to the medical and scientific communities.

Table V-3. Stressors Currently Included in the Navy's Medical Surveillance Programs.

Chemical Stressors

2-Acetylaminofluorene
Acrylamide
Acrylonitrile (Vinyl Cyanide)
Allyl Chloride
4-Aminodiphenyl
Anesthetic Gases
Animal Associated Disease
Antimony
Antineoplastic Drugs
Arsenic
Asbestos
Benzene
Benzidine
Beryllium
Blood and/or Body Fluids
Boron Trifluoride
Cadmium
Carbon Black
Carbon Disulfide
Carbon Monoxide
Carbon Tetrachloride
Chloroform
bis Chloromethyl ether
beta-Chloroprene
Chromic Acid/Chromium VI
Coal Tar Pitch Volatiles/
Polycyclic Aromatic Hydrocarbons (PAH)
Cobalt
Cresol
1,2-Dibromo-3-Chloropropane (DBCP)
3,3'-Dichlorobenzidine
4-Dimethylaminoazobenzene
Dinitro-ortho-Cresol
Dioxane
Epichlorhydrin
Ethoxy and Methoxy Ethanol
Ethylene Dibromibe (EDB)
Ethylene Dichloride
Ethylene Oxide
Ethyleneimine
Fluorides (Inorganic)
Formaldehyde
Glycidyl Ethers
Herbicides
Hydrazines
Hydrogen Cyanide/Cyanide Salts
Hydrogen Sulfide
Hydroquinone (Dihydroxy Benzene)
Isocyanates
Lead (Inorganic)
Machine Oil Mist/Cutting Fluids
Manganese Oxide Fumes
Manmade Mineral Fibers
Mercury
Methyl Bromide
Methyl Chloromethyl Ether
4,4'-Methylene *bis* (2-Chloroaniline) (MOCA)
Methylene Chloride (Dichloromethane)
Methylene Dianiline
alpha-Naphthylamine
beta-Naphthylamine
Nickel (Inorganic)
Nickel Carbonyl
Nitrogen Oxides
4-Nitrobiphenyl
Nitroglycerine
N-Nitrosodimethylamine
Organophosphate/Carbamate Compounds
Organotin Compounds
OTTO Fuel/Other Alkyl Nitrate Propellants
Polychlorinated Biphenyls (PCB)
beta-Propiolactone
Silica (Crystalline)
Styrene
Sulfur Dioxide
1,1,2,2-Tetrachloroethane
Tetrachloroethylene (Perchloroethylene)
Tetryl
Toluene
o-Toluidine
1,1,1-Trichloroethane (Methylchloroform)
Triclhoroethylene
Triorthocresylphosphate (TOCP)
Tungsten and Cemented Tungsten Carbide
Vanadium
Vinyl Chloride
Xylene

Physical Stressors

Cold
Heat
Noise
Radiation — Ionizing
Radiation — Laser
Radiation — Radiofrequency and Microwave
Segmental Vibration
Sight Conservation
Whole Body Vibration

Mixed Exposures

Acid/Alkali (pH < 4.0 or > 11.0)
Metal Fumes
Mixed Solvents

POPULATIONS AT INCREASED RISK

Since the Navy population represents, as indicated previously, a young, healthy, physically fit group of individuals, sensitive subgroups in the Navy do not exist to the extent represented in the general population. There are, however, gender-specific exceptions to that observation in the case of reproductive toxicants and developmental/teratogenic agents. In the case of known male reproductive toxicants, such as dibromochloropropane (DBCP), permissible workplace exposure levels should take the testicular effects into consideration. In the case of substances for which male reproductive toxicity has not been adequately tested, an adjustment in the uncertainty factor may be warranted to provide an adequate measure of protection from such effects. Any such compensatory allowance, however, would have to be considered on a chemical-specific basis, and should not be a routine adjustment for all substances.

A second important consideration is that the overwhelming majority of occupationally exposed females are of child-bearing age. This presents somewhat of a dilemma, since not all pregnancies are planned, and women are not always aware of pregnancy during the first trimester. Due to the impracticality of setting different workplace standards for men and women, and due to the increasing incorporation of women into all areas and subspecialties within the Navy, a workplace standard or action level must provide an adequate level of protection to prevent against potential developmental toxicity. In the case of known developmental toxicants or teratogens, this would certainly be taken into consideration in the establishment of a permissible work-place exposure level. In those cases in which adequate testing for such effects has not been conducted, however, an adjustment in the uncertainty factor for that chemical/substance might warrant consideration. In the case of known developmental/ teratogenic agents, a female worker should be removed from exposure immediately upon learning of the pregnancy. The continued exposure of lactating women should also be considered on a case-by-base, chemical-by-chemical basis.

UNCERTAINTY FACTORS

The factors used to express the uncertainties inherent in the assumptions and data base used to calculate a safe occupational level for Navy uniformed workers might also be somewhat different than those used by EPA and ATSDR for the general population. For the typical Navy occupational risk assessment scenario, some of the uncertainty factors proposed by Barnes and Dourson (1988) apply, while others do not. An appropriate Navy uncertainty factor scheme is presented in Table V-4. The uncertainty factors for interspecies extrapolation would remain, as would the factors for data base

deficiency and use of a lowest-observed-adverse-effect level (LOAEL) for health risk assessment derivation. In a Navy occupational paradigm, however, the use of an additional factor to account for protection of particularly sensitive individuals within the human population (i.e., intraspecies variability) would not usually be considered necessary in many scenarios, since the exposed population in this case represents a particularly healthy, adult, drug-free, physically fit group from which less fit or medically sensitive individuals have been screened, and therefore compositionally different from the more heterogenous general population. In certain cases, however, an additional factor, such as the "2" used by NRC in the establishment of SPEGLs, may be appropriate. Such a determination would have to be made on a chemical-by-chemical/case-by-case basis.

Table V-4. Suggested Uncertainty Factors for Navy Occupational Exposures.*

Factor	Reason
1–10	Interspecies variability (may be < 10 if animal species known to be equally or more sensitive than humans for the chemical or effect under study).
3–10	Use of a LOAEL for risk derivation.
3–10	For severe data base deficiencies if primary effect uncertain.
1–3	Intraspecies variability in sensitivity/susceptibility.
0.1–10	Modifying factor to account for areas not considered in above uncertainty factors (e.g., small sample size in study).

*Subchronic to chronic extrapolation is not considered appropriate for Navy occupational exposures.

A second uncertainty factor routinely employed by EPA, but which would not be necessary or appropriate for Navy occupational assessments, is the factor applied to adjust for the use of a subchronic study, or studies, for a chronic exposure risk estimate (i.e., subchronic to chronic exposure uncertainty factor). Since the typical Navy exposure is both intermittent in nature and of considerably shorter duration than the 30 years assumed by EPA for chronic non-carcinogenic exposure, the Navy occupational exposure is considered to be subchronic in duration, and an uncertainty factor for this area is not needed.

The quantitative definition of subchronic as used by the EPA is approximately 10% of the lifetime of an organism (EPA, 1980). In the case of humans, this would be approximately 7 years (EPA Superfund program uses 2 weeks to 7 years), and would be expected to roughly approximate a 90-day rodent study (EPA, 1989). The ATSDR does not use the same definition or categoric nomenclature as

EPA. Rather, ATSDR breaks human exposure into 3 categories, namely (1) acute (up to 14 days), (2) intermediate (15–364 days), and (3) chronic (365 days or greater) (ATSDR, 1991). In their paradigm, ATSDR estimates duration-specific health guidance values for only those durations supported by human or laboratory animal studies of comparable length. This eliminates the uncertainty factor for extrapolation of subchronic to chronic exposures. In actuality, the body of toxicologic literature taken collectively suggests that a 90-day (subchronic) rodent study would be of sufficient duration to allow for manifestation of all effects that would be seen in a longer, chronic exposure scenario for the vast majority of chemicals. While chronic studies are not generally used for subchronic assessments, they certainly could be used (without application of an additional uncertainty factor) if a chronic study is the only information available. [The corresponding categorization of acute, sub-chronic/intermediate, intermittent, and chronic for description of Navy workplace exposures has been proposed earlier in this paper.]

Modifying factors (additional uncertainty factors) to account for considerations not represented by any of the other uncertainty factors may also be applied to Navy health risk assessments where appropriate. Such factors range from < 1 to 10 and may be used to compensate for additional uncertainty created by a small sample size used in a study, for a steep slope of the dose–response curve, or in consideration of the use of a healthy worker study to determine risk for a potentially more sensitive subpopulation.

The overall (composite) uncertainty factors tend to be loose upper bound estimates (Dourson, 1993) which account for differences in susceptibility between the test and target species and for sensitivity differences within the human population, and reflect the confidence in the final calculated number and data base supporting that number. The resultant effect of this on the final RfDs and RfCs is an estimation of a dose that is likely to be without adverse effects in sensitive individuals for a lifetime of exposure (i.e., a subthreshold dose).

RFD — MRL APPROACH

Once the data base for a particular substance is examined and evaluated as being adequate for risk assessment determination, that study judged to be of sufficient quality and identifying (ideally) both an adverse effect level, the lowest of which is termed a lowest-observed-adverse-effect level (LOAEL), and a no-observed-adverse-effect level (NOAEL) below which no adverse effects have been identified in any other study, is selected as the "critical study" to serve as the basis for determining a protective risk level. If the highest NOAEL approximates an LOAEL from another credible study, an appropriate

adjustment may be applied in the form of an uncertainty factor or modifying factor. The NOAEL (or LOAEL) is then divided by appropriate uncertainty and modifying factors (see Table V-5) to arrive at an upper limit for exposure to the substance in question.

Table V-5. Assumptions and Limitations in the RFD/RFC Process.

Major Assumptions	Major Limitations
1. A population threshold exists. 2. RfD/RfC estimates represent subthreshold doses. 3. Preventing the critical effect protects against all effects.	1. The NOAEL for the critical effect ignores much of the data and often does not distinguish better quality studies from those of lesser quality. 2. Wide dose spacing in the critical study may identify a NOAEL considerably below the actual threshold. 3. Uncertainty factors are imprecise. 4. Risks above these estimates of subthreshold doses are not estimated.

Use of the NOAEL focuses only on the dose representing the NOAEL and does not incorporate information on the slope of the dose-response curve at higher doses at which effects are observed, nor does it distinguish between studies of varying quality. Further, the spacing of the doses in the critical study influence the level identified as the NOAEL.

This conventional NOAEL method suffers from several recognized shortcomings, however. Major assumptions and limitations inherent in the RfD process are presented in Table V-5. From the limitations listed, the artificiality of the NOAEL is apparent. Since the NOAEL and LOAEL used in the risk assessment process come, by convention, from the same scientific study, a wide dose spacing may result in a NOAEL that considerably underestimates the actual threshold for that effect. An even more serious criticism of the NOAEL approach is the fact that the identification of the NOAEL is based upon a statistical test of a hypothesis that the response rate at a particular dose is equal to the response rate of a control group (Dourson, 1993). As the sample size is increased, the test becomes more sensitive, making it more likely that the null hypothesis will be rejected. This tends to push the NOAEL toward a lower value as the size of the sample population is increased.

Nonetheless, the simplistic, threshold-based NOAEL/LOAEL approach to health risk assessment has survived the scrutiny of peer review and years of use by regulatory and other governmental

agencies. It forms the basis of the non-cancer RfD/RfC and MRL approaches to human health risk assessment used by the U.S. EPA and ATSDR, respectively. It is not, however, an absolute prediction of a toxicity threshold, above which toxic effects are likely to occur; rather it represents a level at or below which adverse effects are unlikely to occur during or following the anticipated period of exposure. Similarly, occasional excursions above that level will not necessarily result in toxicity, but the risk of experiencing adverse effects will be expected to increase with increasing frequency and magnitude of excursions above that level.

ALTERNATE APPROACHES

A number of other approaches to the assessment and quantification of health risk are currently in use or under consideration by the risk assessment community. Among these are the benchmark dose (Crump, 1984; 1994), categorical regression (Hertzberg, 1989, 1991), and Bayesian statistical approach (Jarabek and Hasselblad, 1992). These approaches have advantages over existing approaches, but generally have requirements for larger and better data bases. Table V-6 provides a comparison of the strengths and weaknesses of these approaches to chemical risk assessment.

An alternate approach to the RfD — MRL method of risk estimation that has apparent merit is the "benchmark dose" approach (Crump, 1984; Dourson et al., 1985), in which the fitted dose-response curve is used to calculate a dose or concentration for low-incidence responses (e.g., 0.1%, 1%, 5%, etc.). Through application of appropriate uncertainty factors, a no risk estimate/level may then be derived. This approach, which is actually an extension of the more traditional RfD approach, uses the data (NOAEL and LOAEL) from a toxicity study to predict a level at or below which exposure will not cause a significant increase in the incidence of adverse effects. It is not tied to the NOAEL as a rigid value, however, and therefore allows additional flexibility through permitting toxicological judgment in the selection of any point on the curve as an acceptable response incidence.

The process of deriving a benchmark dose consists of modeling the data in the observed range, calculating the upper confidence limit (UCL) on the dose–response curve, and selecting the point on the UCL curve corresponding to a finite percent increase in incidence of an effect. A benchmark dose is calculated for each response that has adequate data. Because the purpose of application of the model is to derive an estimate of dose for a given incidence that is likely to fall within the experimental dose range and does not require extrapolation to estimates far below the experimental dose range, the uncertainty in the estimates is lessened.

Table V-6. Comparison Of Various Approaches To Risk Assessment.

	RfD/RfC-MRL	Benchmark Dose	Categorical Regression	Bayesian Approach
Assumes Existence of Population Threshold	yes	yes	yes	yes
Dependent Upon Experimental Design of Critical Study[1]	yes	no	yes	no
Uses All Data from Critical Study	no	yes	yes	yes
Uses Entire Dose-Response Curve	no	yes	yes	yes
Considers Effects Other Than Critical Effect	no	yes	yes	yes
Combines Data from Other Studies	no[2]	no	yes	yes
Dependent Solely upon Administered Dosages	yes	no	no	no
Uses Uncertainty Factors	yes	yes	yes	yes
Estimates Risk above NOAEL	no	yes	yes	yes
Derives Estimate of Dose within the Experimental Dose Range	no[3]	yes	yes	yes
May Include Estimates below the Experimental Dose Range	yes	yes	no	—
Estimates Risk above Rfd/Rfc	no	yes	yes	—
May Be Used for Assessing the Risk of Multiple Chemicals	no	no	yes	no

[1] sample size, spacing, number of exposure concentrations

[2] The RfD/RfC process considers, but does not combine, data from other studies. Such data are used, however, in considering the strength of the data base in determining the appropriate composite uncertainty factor

[3] The RfD/RfC process selects either the highest NOAEL or, in the absence of a NOAEL, the LOAEL as the basis of the reference dose; while there may be extrapolation from an LOAEL to an estimated NOAEL, there is no interpolation of dose between the effect and no-effect dosages.

[a] It is understood that there are certain situations in which Navy personnel would, by virtue of defense urgencies or operational commitments, actually be exposed to potentially toxic substances for extended (> 8 hr.) daily and prolonged periods of time, thus possibly exceeding the exposures typically seen in civilian occupational settings. Such exposures are not the normal, or more general, type of exposures experienced by Navy men and women in the performance of their daily shipboard or shore-based duties

The first step in the benchmark method is to fit a dose-response model and calculate the lower 95% confidence on the dose associated with 10% (or other) risk (Dourson et al., 1985). This value then serves as a substitute (alternative) to the NOAEL/ LOAEL used in the current process. A number of adjustments are then made to the benchmark dose. First, a species extrapolation adjustment may be applied to account for differences between the test animal and humans. A standard conversion used by

the U.S. EPA for cancer risk extrapolation, but which might be equally appropriate in this situation, is the division of the dose (benchmark in this case) by the cube root of the ratio of human body weight (assumed to be 70 kg for an adult) to animal weight. Another approach to this problem might be to convert the experimental dosage to mg/kg $bw^{0.75}$/day equivalents (Travis et al, 1990). This use of the body weight to the 3/4 power conversion is also a widely accepted method of adjusting the experimental dosage for interspecies variability.

A second adjustment considered in the regulatory toxicology arena is the division of the modeled benchmark dose by an uncertainty factor ranging from 10 to 100. This uncertainty factor is composed of two parts: (1) an interspecies variability conversion factor (typically 10) and (2) an additional factor to account for the fact that the benchmark dose is a "substitute" LOAEL or threshold value, rather than a NOEL or NOAEL (Dourson, 1993). The magnitude of this latter factor would depend on both the seriousness of the effect upon which the benchmark is based and the steepness of the dose-response curve at a given risk of 10%. This particular uncertainty factor would decrease with benchmarks based upon lower incidences of response (i.e., 5%, 1%, etc.), and with increasing steepness of the dose-response curve. If the critical effect is not serious and the slope of the curve is steep, an uncertainty factor of 1 might be appropriate. Conversely, if the effect is serious and the slope is shallow, a factor of 10 would be in order.

Farland and Dourson (1992) examined data from five chemicals listed on EPA's Integrated Risk Information System (IRIS), and estimated benchmark doses using the linearized multistage model used for cancer risk assessment. Modeled benchmark doses were compared with NOAELs and LOAELs for the critical effect used as the basis of RfD derivation. In this comparison, benchmark doses were calculated as the dose associated with 95% UCL on response at 10%. The authors found that the benchmark dose reasonably predicts (within an order of magnitude) the NOAEL dose as found on IRIS for the majority of cases. From that limited analysis, it also appeared that ratios based upon the LOAEL, rather than the NOAEL, is of somewhat better predictive value when the bench-mark is restricted to 10% response.

In the benchmark dose approach, the choice of model may not be critical, since the resultant estimation is within the observed dose range. Therefore, any models that fit the empirical data well are likely to provide similar estimates of the benchmark dose. The advantage of the benchmark dose is that it is considered to be more responsive to the number of animals used at each dosage level and avoids the use of a NOAEL as a rigid value. The benchmark dose uses more of the existing/available toxicity

data, taking into consideration effects other than the "critical" effect, and can be used to obtain a distribution of probabilities around the benchmark. Although more information can be gleaned from its use than with the NOAEL approach, the benchmark approach can result in increased variability in any resulting comparison. This approach may, however, provide some improvement in accuracy over the more simplistic RfD — MRL approach, and appears to be suitable for use in developing Navy occupational risk assessments, if uncertainty factors germane to the exposed population are employed.

Another alternative method of addressing the risk assessment problem is the categorical regression approach (Hertzberg, 1989, 1991). This approach requires a sufficient number of dose groups in each of several categories of effect. In categorical regression, toxicological responses are grouped into four ordered categories of progressive effects: namely, no effects (equivalent to a NOEL), non-adverse effects (comparable to NOAEL concept), mild-to-moderate effects (conceptually comparable to LOAEL), and severe or lethal effects (Equivalent to FEL) (Farland and Dourson, 1992). For any particular dose and duration, this model can be used to estimate the risk of an effect worse than a given category. Since all data are used in the regression in this approach, the specification of an LOAEL is unnecessary.

As described in Farland and Dourson (1992), the categorical regression method assumes the RfD to be the threshold, even though the RfD is by convention considered to represent a sub-threshold dose. All doses above the RfD are therefore believed to be associated with some level of health risk, an assumption that may result in the overestimation of risk at doses approximating the RfD.

The categorical regression model shows the dose-response curve and is of particular utility when risks above the RfD are to be considered for multiple chemicals. This method can be used to distinguish between chemicals posing different health risks, and can identify that chemical representing the more serious risk, when the data for all chemicals under consideration are normalized on a numerical scale. The ability of this method to estimate the risk posed by a single chemical at exposures near the RfD is more limited (Farland and Dourson, 1992), however, due to the inherent assumption in this approach that the RfD actually represents the toxicity threshold.

An additional approach that merits mention is the Bayesian statistical approach of Jarabek and Hasselblad (1992). Although a recent and not yet widely accepted method, this approach, which estimates concentrations of toxicants associated with specific health effects, has distinct advantages in some areas over the current RfD method. The Bayesian statistical approach makes use of the entire response curve from individual studies, while incorporating uncertainty into the response measure. It

may also be used to combine the response data from individual studies on the same chemical, and utilizes statistical combination of similar type data for different responses.

CONCLUSIONS

To address the issue of Navy occupational exposures, the most appropriate predictor of toxic effects is data from other humans exposed to the same substance or substances. Medical surveillance can produce data which, considered with other evidence, will facilitate the determination of standards for limiting exposures in order to prevent injury or disease in the workplace. In the absence of such empirical data, information from laboratory studies using mammalian species appropriate for use as models to predict comparable effects in similarly exposure humans may be used (with the application of appropriate uncertainty factors) as an initial basis for estimation of a safe occupational level. Once applied in a human occupational setting, clinical tests (blood, urine, etc.) should be conducted on a periodic basis (when exposure is established) to ensure the safety of exposed workers and to gather data for possible determination of an appropriate biomarker for future use as an indicator of exposure and/or predictor of toxicity. If known biomarkers exist for exposures to a chemical or chemicals, these should be used to identify significant human exposure or impending toxicity, as appropriate.

Appropriate clinical tests might include gross examination, full serum and whole blood chemistries, urine chemistry, pulmonary function tests (where appropriate), electrocardiogram, and neurological and neurobehavioral testing appropriate for the chemical(s)/substance(s) of concern. A listing of the various neurobehavioral tests available and the optimal conditions for their respective uses are available in Hutchinson et al. (1992).

The results of the aforementioned periodic clinical tests would be forwarded by the command holding the individual's medical record to a central authority, such as the Navy Environmental Health Center (NEHC), for inclusion on a registry specific for the chemical in question. While there would necessarily be administrative cost associated with the maintenance of such a registry system, the value to the Navy would be great. Responsibility for maintaining the registries could be easily fit into the existing Navy medical organizational structure.

In conclusion, it may well be that no single approach represents a "one size fits all" solution to the problem of occupational health risk assessment for Navy personnel. Not all models are suitable for all data and effect scenarios. Statistical issues such as the applicability of the assumptions inherent in an approach and fit of the model to the data must be considered, and judgment is instrumental in the evaluation and interpretation of the data and selection of the most appropriate approach for use in any

given situation. Judicious correlation of any quantitative assessment with existing clinical data and observations as would come from the Navy's medical surveillance program would serve to strengthen any assessment and should be an integral part of any quantitative health assessment. And in the final analysis, it must be borne in mind that (as paraphrased from Farland and Dourson, 1992) the scientific judgment of experts is critical to the defensibility of any quantitative risk evaluation.

The approaches to occupational health risk assessment considered in this paper are intended for consideration for application to exposures of Navy military personnel, only, and are not intended to apply to Department of the Navy (DON) civilian employees and contractors, whose workplace exposure standards are set by OSHA.

SUMMARY

The exposures of Navy personnel to chemicals and other hazardous substances in the workplace occur during the course of their normal daily duties, as well as in certain atypical or emergency situations. While suitable methodological approaches to assessment of the extent of any risk associated with both chronic and emergency situations are currently in existence, the risk resulting from the intermittent exposures experienced by many Navy people may not be adequately addressed by existing methodologies. The RfD/RfC — MRL, benchmark dose, categorical regression, and Bayesian statistical approaches all represent methodologies of merit; each has its strengths, while each has certain limitations, as well. An approach that may be appropriate for Navy use is the benchmark dose, which is a model-derived estimate of the lower confidence limit on the effective dose that produces a certain increase in incidence above control levels, such as 1, 5, or 10%. This approach, with its estimation of toxicity threshold within the experimental dose range, appears to be most worthy of further investigation for Navy use in cases where sufficient substance-specific data are available. Other statistical approaches also have merit for Navy application, particularly in the area of multichemical exposures, and likewise warrant further examination in this context. The rigid adherence of the RfD/MRL approach to an empirical NOAEL and the uncertainty of the precise location of the toxicity threshold with respect to the RfD make these methods perhaps too conservative for Navy use in some situations, given the health and age of the exposed population and the nature of Navy occupational exposures.

A separate, but related and equally important, issue is the use of the currently existing medical surveillance program for monitoring occupationally exposed personnel within the Navy, and the need for channeling that information into a centralized Navy Medical data base. The data base could then be

maintained in a manner similar to the disease registries of ATSDR or similar manner in which the medical information could be stored and categorized, but the identity of exposed individuals protected. Since the Medical Surveillance Program is already in place, all that would be necessary for the establishment and maintenance of a centralized data base is to forward the input from that system into the data base, which could be maintained by NEHC or other component of the Navy Medical Community. The surveillance system, if vigorously maintained and judiciously utilized, could represent a potentially invaluable source of information on real-time exposure of a known population of healthy individuals and serve to further the knowledge and understanding of the effects of exogenous substances in the workplace on human health. The ultimate benefit of such a program would be an enhanced ability to foresee exposure levels or circumstances beyond which individuals should no longer be exposed, thereby ensuring the protection of human health and the ability of Navy personnel to perform at a maximum health and operational level.

REFERENCES

ATSDR. 1991. *Guidelines for the Development of Environmental Health Assessments.* U.S. Department of Health and Human Services, Public Health Service, Agency for Toxic Substances and Disease Registry, Atlanta, GA.

Barnes, D.G. and M.L. Dourson. 1988. Reference dose (RfD): Description and use in health risk assessments. *Regul. Toxicol. Pharmacol.* 8:471–486.

Bruner, R.H., E.R. Kinkead, T.P. O'Neill, C.D. Flemming, D.R. Mattie, C.A. Russell, and H.G. Wall. 1991. The toxicological and oncogenic potential of JP-4 jet fuel vapors in rats and mice: 12-Month intermittent inhalation exposures. *Fund. Appl. Toxicol.*

Conway, T.L., L.K. Trent, and S.W. Conway. 1989. (Physical Readiness and Lifestyle Habits among U.S. Navy Personnel during 1986, 1987, and 1988.)

Crump, K.S. 1984. A new method for determining allowable daily intakes. *Fundam. Appl. Toxicol.* 4:854–871.

Crump, K.S. 1994. Calculation of benchmark doses from continuous data. Pre-print. Submitted to *Risk Analysis* on April 11, 1994.

DoD. 1984. Department of Defense Instruction 6055.1, Change 1 of 4 May 84.

Dourson, M.L. 1993. Overview of alternatives to benchmark dose. Presentation at EPA-ILSI Workshop on Benchmark Dose. Washington, D.C. September 23, 1993.

Dourson, M.L., R. Hertzberg, R. Hartung, and K. Blackburn. 1985. Novel methods for the estimation of acceptable daily intake. *Tox. Ind. Health* 1(4):23–33.

EPA. 1980. Guidelines and Methodology for the Preparation of Health Effects Assessment chapters of the Ambient Water Quality Criteria Documents. United States Environmental Protection Agency. Office of Water Regulations and Standards. Criteria and Standards Division. Washington, D.C.

EPA. 1989. Risk Assessment Guidance for Superfund Volume I Human Health Evaluation Manual (Part A). United States Environmental Protection Agency. Office of Emergency and Remedial Response. Washington, D.C. pp. 6–2.

EPA. 1990. Interim Methods for Development of Inhalation Reference Concentrations. United States Environmental Protection Agency. Office of Research and Development. Washington, D.C. EPA/600 8-90/066A.

Farland, W. and M. Dourson. 1992. Noncancer health endpoints: Approaches to quantitative risk assessment. *In: Comparative Environmental Risk Assessment*, C.R. Cothern, ed., Lewis Publishers, Ann Arbor, MI, 1992.

Flemming, C., O.M. Little, H. Barton, R. Carpenter, J. Gearhart, and H. Clewell. 1993. The effect of body fat and body weight from Navy subpopulations on dose metrics used in risk assessments. Conference on the Risk Assessment Paradigm after Ten Years: Policy and Practices Then, Now, and in the Future. April 5–8, 1993, Wright-Patterson Air Force Base, Dayton, OH.

Hertzberg, R.C. 1989. Fitting a model to categorical response data with application to species extrapolation of toxicity. *Health Phys.* 57:405–409.

Hertzberg, R.C. 1991. Quantitative extrapolation of toxicologic findings. In: *Statistics in Toxicology*. D. Krewski and C. Franklin, eds. Gordon and Breach, New York. pp. 629–652.

Hutchinson, L.J., R.W. Amler, J.A. Lybarger, and W. Chappel. 1992. Neurobehavioral Test Batteries for Use in Environmental Field Studies. U.S. Department of Health and Human Services, Public Health Service, Agency for Toxic Substances and Disease Registry, Atlanta, GA.

Jarabek, A.M. and V. Hasselblad. 1992. Application of a Bayesian statistical approach to response analysis of noncancer toxic effects. *The Toxicologist.* 12(1):98 (no. 305).

NEHC. 1991. Medical Surveillance Procedures Manual and Medical Matrix (4th Edition, dated September 1991). Navy Environmental Health Center, Norfolk, VA.

NMC. 1988. May 2, 1988 letter from the Commander, Naval Medical Command, Washington, D.C. *re* Worker Placement Into Medical Surveillance Programs.

NRC. 1986. Criteria and Methods for Preparing Emergency Exposure Guidance Level (EEGL), Short-Term Public Emergency Guidance Level (SPEGL), and Continuous Exposure Guidance Level (CEGL) Documents. Committee on Toxicology, Board on Environmental Studies and Toxicology, Commission on Life Sciences, National Research Council. National Academy Press, Washington, D.C.

NRC. 1993. Guidelines for Developing Community Emergency Exposure Levels for Hazardous Substances. Subcommittee on Guidelines for Developing Community Emergency Exposure Levels (CEELs) for Hazardous Substances, Committee on Toxicology, Board on Environmental Studies

and Toxicology, Commission on Life Sciences, National Research Council. National Academy Press, Washington, D.C.

OPNAV. 1983. Chief of Naval Operations (OPNAV) Instruction 5100.23B of 31 August 1983.

Travis, C.C., R.K. White, and R.C. Ward. 1990. Interspecies extrapolation of pharmacokinetics. *J. Theor. Biol.* 142: 285–304.

U.S. Air Force. 1974. Chronic inhalation toxicity of JP-4 jet fuel. In: Toxic hazards research unit annual technical report: 1974. Report no. AMRL-TR-74-78. Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory, Aerospace Division, Air Force Systems Command. pp. 5–26.

U.S. Air Force. 1984. Ninety-day continuous inhalation exposure to petroleum JP-4 jet fuel. In: Toxic hazards research unit annual technical report: 1984. Report no. AMRL-TR-84-001. Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory, Aerospace Medical Division, Air Force Systems Command. Document no. AD-A147857/7, 46–62.

U.S. Census Bureau. 1993. U.S. Bureau of Census Statistical Abstract of the United States, 113th Ed.

TEMPORAL ASPECTS OF RISK CHARACTERIZATION

M.J. Goddard[1], D.J. Murdoch[2], D. Krewski[3]

[1]Environmental Health Centre
Tunney's Pasture,Ottawa, Ontario, Canada, K1A 0L2
[2]Department of Mathematics and Statistics
Queens University, Kingston, Ontario, Canada, K7L 3N6
[3]Environmental Health Centre
Tunney's Pasture, Ottawa, Ontario, Canada, K1A 0L2

ABSTRACT

Although many laboratory studies are conducted with the dose level held constant for the duration of the experimental period, human exposure to toxic substances can vary widely over time. In this article we discuss methods for toxicologic risk assessment with time-dependent exposure patterns. In particular, methods for carcinogenic risk assessment under both the multi-stage and two-stage models are reviewed. It is shown that the use of a "Lifetime Average Daily Dose" or "LADD" can overestimate or underestimate the actual lifetime risk associated with time-dependent exposure patterns. It is also shown that there exists a "Lifetime Equivalent Constant Dose" or "LECD" that leads to the same lifetime risk as the actual time dependent exposure pattern. The ratio C = LADD/LECD thus provides a measure of accuracy of the LADD. Several applications are summarized herein to illustrate the application of currently available risk assessment methods for time-dependent

Key Words: Time Dependent Exposure Patterns, Multi-Stage Model, Two-State Model, Pesticides, Benzo[*a*]pyrene, Radon, Tobacco, Interaction.

Introduction

The toxicologic basis for risk assessment has undergone some noteworthy changes in recent years, including a shift in emphasis from observational toxicology to biological and mechanistic studies (Goddard and Krewski, 1995). One benefit of this new focus is the opportunity to examine the impact of assumptions that have often been made in toxicologic risk assessment. In this article, we focus on risk assessment problems in which the level of exposure to the agent of interest varies over time. Risk assessment methods developed for use with toxicologic and epidemiological data involving time-dependent exposure patterns are described, and illustrated using data from a series of studies in which the level of exposure is not constant.

In the absence of risk assessment methods for time-dependent exposures a constant dose is used. Typically, exposure profiles characterized by the duration and intensity of exposure are noted, and an average dose calculated over the time period of interest, often an entire lifetime. Estimates of risk are then calculated under the assumption that this average dose would approximate the cumulative lifetime risk associated with the actual time-dependent exposure pattern. The U.S. EPA adopted this as a guideline in 1986 (U.S. Environmental Protection Agency, 1986):

> "Unless there is evidence to the contrary ... the cumulative dose received over a lifetime, expressed as an average daily exposure pro-rated over a lifetime, is recommended as an appropriate measure of exposure to a carcinogen. That is, the assumption is made that a high dose of a carcinogen received over a short period of time is equivalent to a corresponding low dose spread over a lifetime."

There is now a substantial literature on risk assessment methods for use with time-dependent exposure patterns. Whittemore and Keller (1978) presented a detailed analysis of the Armitage-Doll multi-stage model of carcinogenesis, allowing for the effects of intermittent exposure. Subsequently, Day and Brown (1980) extended this work to examine the effects of allowing different stages to be dose-dependent. Crump and Howe (1984) published detailed formulae for the multi-stage model in which one or more stages were dose dependent, and used these results to estimate the potential cancer risks of exposure to ethylene dibromide. Morrison (1987) showed that for some models, detailed exposure histories are needed, and that the use of simple time-averaged dose approximations may lead to errors of several orders of magnitude for both the multi-stage and clonal two-stage model estimates. Kodell, Gaylor, and Chen (1987) conducted a theoretical investigation of the effects of time-dependent exposure under the multi-stage model, and established bounds on the degree of error associated with an average daily dose in place of the actual time-dependent exposure pattern. Chen, Kodell, and Gaylor (1988) and Murdoch and Krewski (1988), subsequently extended this work to the two-stage birth-death-mutation model discussed by Moolgavkar and Lubeck (1990). This biologically based two-stage model of carcinogenesis incorporates important aspects of neoplastic transformation such as tissue growth and cell kinetics not included in the Armitage-Doll model.

In this paper, we provide a brief overview of the multi-stage and two-stage models of carcinogenesis (Section 2). In Section 3, we introduce the concept of a lifetime equivalent constant dose or (LECD) as a basis for comparison with the lifetime average daily dose or LADD. The LECD leads to the same cumulative risk as the actual time-dependent exposure profile under consideration,

and provides a benchmark against which to evaluate the accuracy of risk estimates based on the LADD. Four applications of risk assessment methods for time-dependent exposure patterns are presented in Section 4. The first involves the establishment of maximum allowable air concentrations of volatile organic compounds for astronauts spending a limited period of time in a manned space station. The second addresses the risks associated with dietary exposures to pesticides, which can vary substantially over time due to changes in food consumption patterns. The third example is based on a laboratory experiment in which groups of mice were exposed to different doses of benzo[a]pyrene with widely varying exposure profiles. The fourth application considers the effects of joint exposure of uranium miners to exposure to two agents, radon and cigarette smoke. Our conclusions are presented in Section 5.

Models of Carcinogenesis

The Armitage-Doll multi-stage model (Armitage and Doll, 1957; Armitage and Doll, 1961) has a long history of application in carcinogenic risk assessment. The basic idea underlying this model is that normal cells undergo an ordered sequence of transitions progressing through a series of intermediate stages to become cancerous. The rates of transitions are modeled with an intensity parameter λ for each stage. Note that in this formulation the transition intensities do not depend directly on the subject's age, and that they are constant over the course of a lifetime. The effects of exposure to a carcinogenic substance are introduced by letting λ_i depend on the dose d. This dependence is usually represented in terms of a linear function $\lambda_i = a_i + b_i d$. Here, a_i represents a background transformation rate and b_i reflects the effects of exposure to the agent of interest. In a k stage model, the product of the λs defines the cumulative hazard function $H(t; d)$:

$$H(t;d) = \frac{t^k}{k!}\prod_{i=1}^{k}(a_i + b_i d). \tag{1}$$

The probability of a tumor at time t and dose d is then given

$$P(t;d) = 1 - \exp\{-H(t;d)\}. \tag{2}$$

This model is capable of describing many datasets, particularly age-specific human cancer incidence curves which are linear on a log-log scale. However, this model does not fit all datasets equally well, especially datasets involving childhood cancers.

Currently there is considerable interest in a modified form of the model, called the two stage birth-death-mutation or TSM (Moolgavkar, 1986; Moolgavkar and Luebeck, 1990). This approach restricts the multi-stage model to two stages, but it allows for the kinetics of normal and intermediate cells. Although approximate solutions for the two-stage model have been applied in practice [see (Krewski, Goddard and Zielinski, 1992) for a recent review], the exact solution for the TSM is somewhat complex (Denes and Krewski, 1994). The essential mathematical details can be found in Moolgavkar and Luebeck (1990). In practice, appreciable flexibility is gained in describing both toxicologic and epidemiological data by allowing for tissue growth and cell kinetics in addition to two critical mutations.

In this model, intermediate cells are formed from normal cells that have sustained the first mutation; a malignant cancer cell is produced from an intermediate cell that sustains the second critical mutation. The rate of transformation of normal to intermediate cells is modeled by λ_1 and the rate from intermediate to cancerous cells is modeled by λ_2. Normal cells are regenerated at the rate of α_1 and die at the rate of β_1, the corresponding birth and death rates for intermediate cells are α_2 and β_2 respectively.

The model affords a useful framework which clearly distinguishes mutation from cellular dynamics. Operational terms like "initiation" and "promotion" have natural interpretation: an initiator is defined as an agent which increases the first mutation rate; a promotor increased the net birth rate of the intermediate cell population. In both cases, the pool of cells at risk of malignant transformation is increased, although the mechanisms by which this is accomplished are quite different. Net promotional activity is modeled by $\delta = \alpha_2 - \beta_2$, the net growth rate of intermediate cells. Note that no effect on cellular transformation has been assumed.

Temporal Patterns of Exposure

Most of the risk assessment applications to date based on the models discussed above has involved the simplifying assumption that exposure to a factor has been at a constant level throughout the life of the individual. In this section, we summarize a method for incorporating varying exposure patterns in the multi-stage and the TSM models developed by Murdoch *et al.* (Murdoch and Krewski, 1988; Murdoch, Krewski and Wargo, 1992).

It is convenient to express the probability of a response at time t under the multi-stage model in terms of the cumulative hazard function $H(t; d)$ (equation 2). Let $d(t)$ be a function that describes how

dose varies with time. Assuming that dose affects the transition intensity function for one stage in the model, say stage r, then the cumulative hazard $H_r(t)$ is a linear function of d_r^*:

$$H_r(t)=a(t)+b_r(t)d_r^*(t) \tag{3}$$

where

$$d_r^*=\frac{1}{t}\int_0^t d(u)tw(u;r,\ k,t)du \tag{4}$$

Introducing

$$R(u)=tw(u;r,k,t) \tag{5}$$

with

$$\frac{1}{t}\int_0^t R(u)du = 1 \tag{6}$$

we may write

$$d_r^*=\frac{1}{t}\int_0^t d(u)R(u)du \tag{7}$$

The value d_r^* represents that level of exposure which, if administered at a constant rate from time 0 to time t, would produce the same cumulative hazard at time t as the dose $d(u)$ which varies in time. This value (d_r^*) will be referred to as the life time equivalent constant dose or LECD. The term $w(u;\ r,\ k,\ t)$ is a weighting function that can be computed using statistical expectations of order statistics (cf. [Murdoch and Krewski, 1988]). $R(u)$ is weighting function that describes the relative effectiveness of dosing at different points in time. This relative effectiveness function depends both on the number of overall stages and that stage which is dose dependent.

Figure V-1 shows the relative effectiveness function $R(u)$ for a six stage model with different stages dose-dependent. Most of the weight occurs near the start of life if an early stage is affected. When a later stage is assumed to be dose dependent, the greatest weighting occurs at later ages.

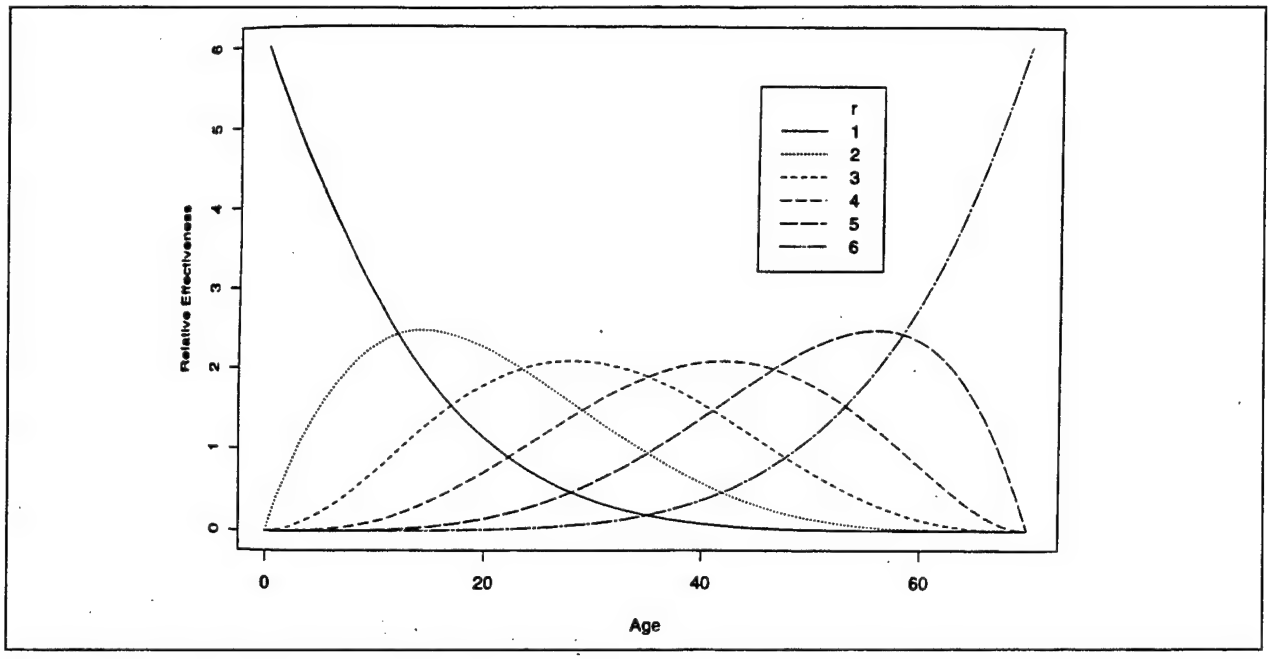


Figure V-1. Weighting Functions for Six Stage Model.

Let $\bar{d}$ represent the prorated average daily exposure described by the EPA guidelines (U.S. Environmental Protection Agency, 1986):

$$\bar{d}=\frac{1}{t}\int_0^t d(u)(du). \tag{8}$$

This represents the cumulative dose up to time t divided by the length t of the time period of interest, giving an unweighted average dose. If t corresponds to an expected lifetime, we will refer to $\bar{d}$ as the lifetime average daily dose or LADD. Note that now allowance is made for the time at which exposure occurred when computing the LADD:

Let C represent the ratio of the LECS to the LADD:

$$C=\frac{d_r^*}{\bar{d}}=\frac{LECD}{LADD}. \tag{9}$$

Values of C less than one represent cases in which the LECD is smaller than the LADD; in such case, the risk associated with the LADD $\bar{d}$ will exceed the actual risk based on the LECD. On the other hand, if C > 1, then the risk associated with $\bar{d}$ will be less that the actual risk.

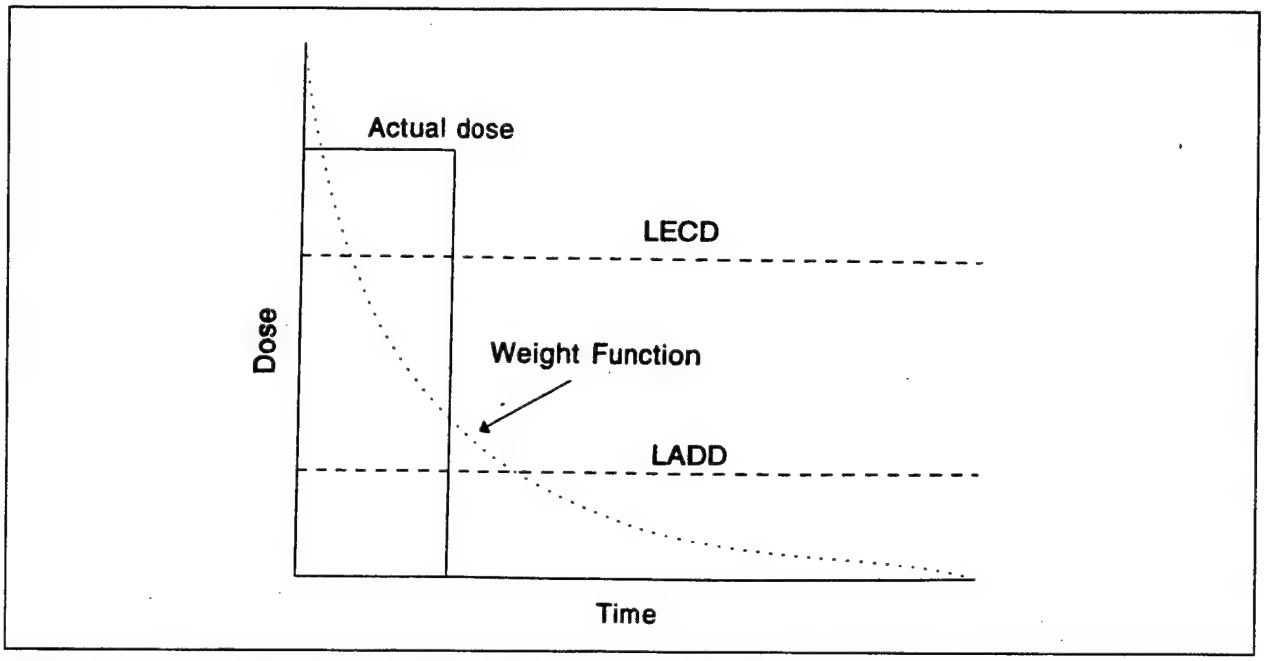


Figure V-2. Example Showing the LECD and the LADD.

To illustrate, Figure V-2 depicts a situation in which exposure occurs only during the first quarter of a lifespan. The solid line represents the actual time dependent exposure profile, $d(t)$, and the lower dashed line corresponds to the LADD ($\bar{d}$). The dotted line shows the relative effectiveness function $R(t)$ for a 6 stage model where the first stage is dose dependent. The upper dashed line shows the lifetime equivalent constant dose d_1^*. Here, the LADD is $0.25d$ and the LECD works out to $0.82d$; the ratio, C, is therefore 3.28. The use of the LADD would therefore underestimate the actual risk by about this factor.

An LECD can also be determined for the TSM model (Murdoch and Krewski, 1988). It is possible to obtain simple expressions for the hazard function for this model with the first and second stage transition intensity functions λ_1 and λ_2. As in the multi-stage model, λ_1 and λ_2 are assumed to be linearly dependent d. If only one stage is dose dependent, then the LECD can be expressed as a weighted average of the time dependent dose similar to that developed for the multi-stage model.

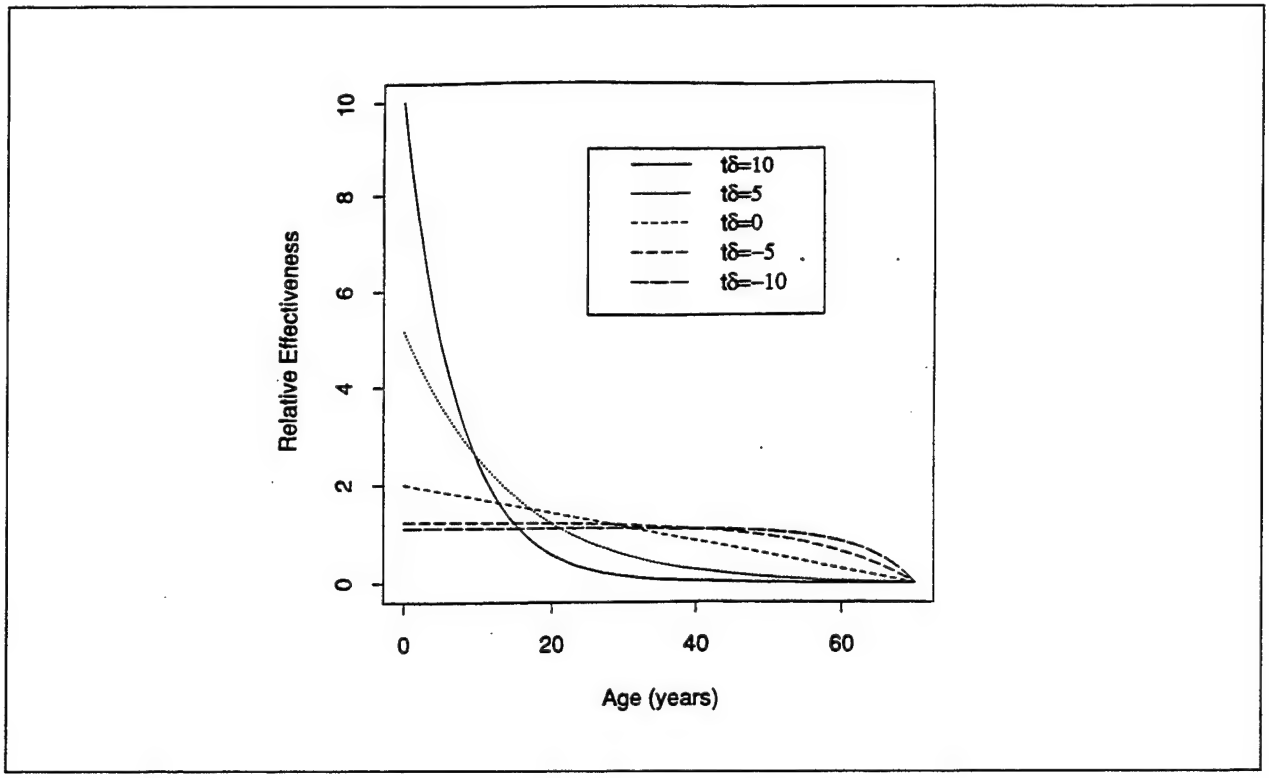


Figure V-3. Weighting Functions *[R(u)]* for the TSM Model.

Figure V-3 shows the relative effectiveness functions for the TSM model where only the transformation between normal and intermediate cells (the first stage) is dose dependent. The different curves are characterized by different values of $t\delta$, which represents the net birth rate of initiated cells over an interval of length t. Note that because only the first stage is considered to be dose-dependent, earlier exposures are more important than later exposures. The curves for large, negative values of $t\delta$ model initiated cells that die rapidly, (i.e., are shortlived) and it is seen that the relative effectiveness in these situations is almost constant. Larger positive values of $t\delta$ yield relative effectiveness functions which vary much more with time.

Examples

Spacecraft Exposures

One use of C is to measure the extent of over or underestimation of risks through the use of the simple dose average (LADD) rather than a weighted equivalent dose (LECD). To illustrate this application in a practical context, consider first the need to establish air quality guidelines for astronauts spending a limited period of time in a closed environment provided by the manned space station planned by NASA (National Research council, 1992). These individuals will breath recycled

air in an artificial and self-contained environment for up to 180 days. Values of the lifetime equivalent constant dose and the lifetime average daily dose were calculated for several scenarios. The age of the astronaut at the start of exposure was taken as 25, 30, 35, 40, and 45 years. The duration of exposure was assumed to be 1, 30, and 180 days. For the multi-stage model up to 6 stages were assumed with only one stage dose dependent. All admissible stages were individually considered to be that which is dose dependent.

Table V-7. Values of *C* for a 45-Year-Old Astronaut Calculated Using a Six-Stage Model for a 30-Day Duration.

Age. (years)	Stage Affected *r*	Number of Stages *k*					
		1	2	3	4	5	6
45	1	1.00	0.71	0.38	0.18	0.08	0.03
	2		1.29	1.38	0.98	0.58	0.31
	3			1.24	1.77	1.58	1.13
	4				1.07	1.90	2.03
	5					0.86	1.83
	6						0.66

Table V-7 gives the values of *C* for a 45-year-old astronaut exposed for 30 days. This table shows one of the widest ranges for values of *C* among all the scenarios considered. The lowest value, 0.03, corresponds to a relatively old astronaut with the first stage of the multi-stage model being dose-dependent. This shows that the simple dose average can overestimate the equivalent constant dose by almost two orders of magnitude. When the fourth stage is assumed to be dose dependent, we observe a value of 2.03 for *C*, which would correspond to overestimation of risk by about a factor of two-fold.

Values of *C* under various scenarios varied by only at most 5% among the 1, 30, and 180 day exposures. The implication is that the ratio of the LECD to the LADD is relatively constant for these exposures. This observation derives from the fact that 180 days is a very small part of a 70 year lifespan, and that during that relatively short period of the astronaut's lifetime, the relative effectiveness function is approximately constant.

Similar results can be obtained for the TSM (National Research Council, 1992). Under this model, the values of *C* ranged from 0.016 to 1.4 under comparable assumptions (Murdoch, Krewski, and Wargo, 1992).

Exposure to Pesticides During Infancy

As individuals mature, dietary patterns typically vary. Risks which may be associated with toxicants combined with food are therefore expected to vary as well. Consider the problem of determining risks among children exposed to pesticides associated with apple juice (National Research Council, 1993). Elevated exposures in youth may be of greater toxicologic significance than exposures at later ages. One reason may be that in youth there is greater cellular growth and development which may render tissues more susceptible than those in adults. Infants and children have different food consumption patterns than adults. For example, the National Food Consumption Survey conducted by the U.S. Department of Agriculture (U.S. Department of Agriculture, 1978) showed that ingestion of apple juice on a per kilogram body weight per day basis peaked in younger children and declined as an individual aged. Exposure to any pesticides associated with these foods is clearly not constant across the individual's lifetime.

Table V-8. Values of C for Infants Exposed to Pesticides Calculated Using a Six-Stage Model.

Food	Stage Affected *r*	Number of Stages *k*					
		1	2	3	4	5	6
Apple Juice	1	1.00	1.64	2.23	2.77	3.27	3.74
	2		0.36	0.48	0.61	0.76	0.93
	3			0.30	0.34	0.38	0.43
	4				0.28	0.31	0.34
	5					0.27	0.30
	6						0.27

Values of *C* for apple juice consumption are given in Table V-8. (Murdoch, Krewski, and Wargo, 1992). The largest value of *C* is larger than in the previous example: this is due primarily to the strong age-dependence of the exposure. Note particularly the effect in the multi-stage model when the earlier stages are dose dependent, compared to the fit when later stages are assumed to be dose dependent. Using the TSM model, *C* varied from 0.25 to 5.25.

Benzo[a]pyrene

While intermittent exposures are typical of human exposures, they are not routinely used in laboratory experiments. One laboratory experiment with highly varying dose patterns was reported by Neal and Rigdon (1967). This experiment was designed to explore the dose-response relationship

between gastric tumors and oral exposure to benzo[*a*]pyrene in mice. Dose levels varied form 0.001 to 5 mg/kg, with periods of dosing ranging from 2 to 152 days.

Krewski and Murdoch (1990) analyze these data using both a two-stage version of the Armitage-Doll model and an approximate form of the two-stage birth-death-mutation model. Table V-9 presents the dosing regimens as well as the observed and the predicted numbers of gastric tumors. The predictions for both models did not fit the observed data in an entirely satisfactory manner. Reasons for this include the use of the approximate form of the two-stage birth-death-mutation model, known to be inaccurate for the high response rates encountered in this study, and the noteworthy variability of the results, particularly between the response rates in groups 7 and 8.

In this analysis two stages rather than one are assumed to be dose-dependent. The cumulative hazard function is more complicated than (3), depending on d_1^* and d_2^* as well as d_{12}^*, a two-dimensional analogue of d_r^*. From the coefficients of these terms, it is possible to estimate the ratios b_1/a_1 and b_2/a_2, the ratios of the slope to the constant terms in the transformation rates λ_1 and λ_2. Values of b_1/a_1 and b_2/a_2 were approximately 5.8×10^{10} and 1.5×10^{-1} respectively. This is taken as evidence that the carcinogenic effects of benzo[*a*]pyrene are stronger for the first stage (initiation) of the carcinogenic process. Some evidence of a positive effect of proliferation $\delta = \alpha_2 - \beta_2$ on initiated cells was also noted.

Values of C are also given in Table V-9, where it is seen that for the Armitage-Doll model, the fraction ranges from 0.49 to 1.36. For the two–stage birth–death–mutation model, these values range from 0.07 to 1.12. The lowest values of C under the TSM occurred for dose patterns in which B[a]p was administered late in life.

Table V-9. Predicted and Observed Forestomach Tumor Rates in Mice Exposed to Benzo[*a*]pyrene.

Dose Group	Dose[1] *d*	Exposure[2] Days	Sacrifice Day	Number Exposed	Animals with Tumors			Values of *C*	
					Observed	A–D[3]	TSM[4]	A–D	TSM
1	0	0–300	300	289	0	0.0	0.0	–	–
2	0.001	30–140	140	25	0	0.2	0.2	0.79	0.50
3	0.01	30–140	140	24	0	1.9	1.4	0.79	0.50
4	0.02	116–226	226	23	1	3.6	2.7	0.49	0.07
5	0.03	50–160	160	37	0	8.3	6.3	0.69	0.32
6	0.04	67–177	177	40	1	11.4	8.8	0.62	0.22
7	0.045	51–161	163	40	4	13.0	10.4	0.70	0.32
8	0.05	20–172	172	34	24	18.8	23.8	0.88	0.64
9	0.1	22–132	132	23	19	13.3	11.1	0.83	0.60
10	0.25	19–137	137	73	66	67.0	66.6	0.86	0.65
11	0.25	49–50	155	10	0	0.4	0.4	1.36	1.12
12	0.25	56–58	162	9	1	0.6	0.8	1.30	0.94
13	0.25	49–53	155	10	1	1.3	1.6	1.34	1.08
14	0.25	62–67	168	9	4	1.5	1.7	1.23	0.79
15	0.25	49–56	155	10	3	2.2	2.5	1.32	1.03
16	0.25	91–121	198	26	26	16.1	15.5	0.93	0.29
17	0.1	74–81	182	10	0	1.0	1.1	1.15	0.59
18	0.1	48–78	156	18	12	5.8	5.6	1.19	0.79
19	5	139–140	252	33	17	17.9	21.9	0.89	0.15

[1] Doses in mg Benzo[*a*]Pyrene per *g* food
[2] In some cases, where the start and end points are known to lie within a certain interval, the midpoint of the interval is shown.
[3] Armitage–Doll Multistage model.
[4] Two stage birth–death–mutation model.

Radon and Cigarette Smoke

All of the previous applications have involved exposure to a single (time dependent) risk factor. Although this situation is the easiest to deal with using the methods discussed here, it does not reflect the human environment in which multiple exposures are common. Studies of mixtures of agents are receiving increasing attention (National Research council, 1988). Although algebraically involved, there is little theoretical difficulty in generalizing the intermittent exposure considerations described above to multiple agents.

The joint effect of radon and cigarette smoke in the induction of lung cancer among uranium miners in the Colorado Plateau was considered by Moolgavkar *et al.*, (1993) within the context of the two-stage model. Figure V-4 summarizes seven configurations of joint exposure to radon and tobacco

smoke experienced by the miners. The first row of the figure represents miners who were exposed to radon but who did not smoke, whereas the second shows miners who stopped smoking before they were exposed to radon. The exposure times for each miner were considered in the analysis; here, the size of the bars and rectangle in Figure V-4 are not to scale. The complexity of these exposure patterns precludes the derivation of simple algebraic formulae for the cumulative hazard, such as that given in (1). Statistical analysis of data using such models is based on intricate computer algorithms used to fit the model to the available data.

Moolgavkar *et al.*, (1993) based their analysis on both individual records of exposure from Colorado miners and from aggregate information from the British doctors' data. The number of miners at risk in each dose-pattern category, along with the observed and expected incident cases are shown in Figure V-4 as well.

Studies exploring mixtures of agents frequently allow one to explore the effects of the individual agents as well as any possible synergy the agents confer on each other on the different stages. Where intermittent exposures occur, there is also the question of whether the order of exposure is important. Using the two-stage birth-death-mutation model with deterministic growth of normal cells, Moolgavkar *et al.*, (1993) were able to exploit the time dependent exposures to the different agents and conclude the following: (1) Both radon and cigarette smoke affect the first mutation rate as well as the birth-death dynamics of intermediate cells, but not the second mutation rate; (2) The order of exposures to the different agents had only a modest effect on risk; (3) Age-specific relative risks associated with joint exposures are supra-additive (that is, greater than the sum of the individual effects) but sub-multiplicative (that is, less than the product of the two effects).

Discussion

In this paper, we have discussed methods for evaluating the effects of varying temporal patterns of exposure in risk assessment. The methods described here are primarily applicable in the area of cancer risk assessment, and are based on the Armitage-Doll multi-stage model or the two-stage birth-death-mutation model.

It was shown how a lifetime equivalent constant dose d_r^* can be determined for both models using the concept of a relative effectiveness function to indicate the importance of exposures occurring at different points in time. The ratio of the LECD to the lifetime average daily dose or LADD commonly used in risk assessment application with time-dependent exposure was used to evaluate overestimation or underestimation of risk based on the LADD.

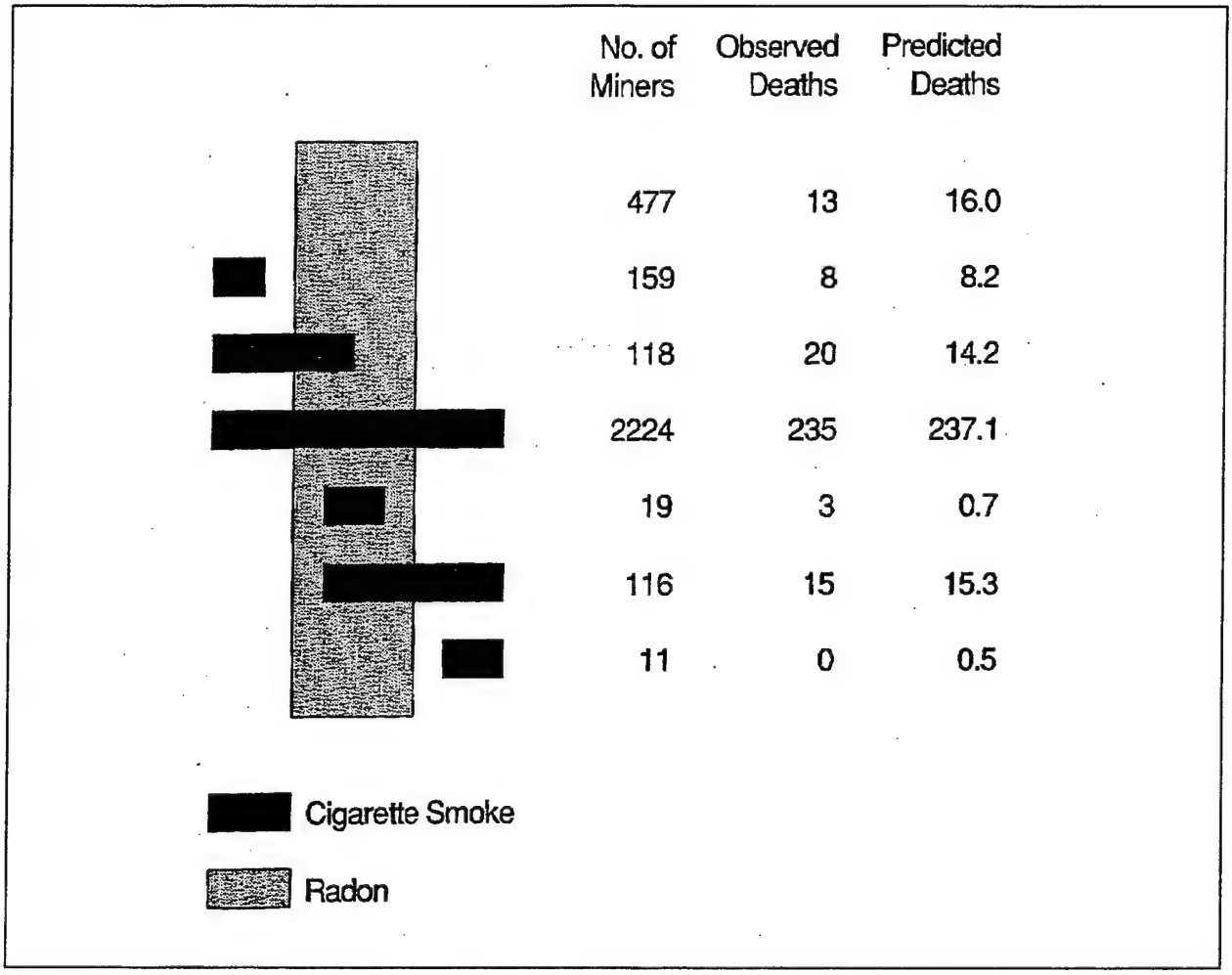


Figure V-4. Combinations of Exposures to Radon and Cigarette Smoking. (Bar lengths are not scaled to reflect observed exposure duration.)

Herein, we have focused on a comparison of the LADD and the LECD using the ratio C = LADD/LECD. One limitation in using an LECD approach to estimate risks is the dependence of the LECD on the weighting function $R(u)$ which itself depends on the model and the stage in carcinogenesis the agent is believed to affect. In many cases, the stage most affected will not be clearly known and various scenarios will need to be considered.

Based on applications involving astronauts exposed to volatile organic materials in a closed station environment and pesticide residues in the diets of infants and children, the values of C were generally not much greater than unity, indicating a limited degree underestimation of risk using the LADD. However, values of C much lower than unity were also noted, raising the possibility of substantial overestimation of risk. Two additional examples focused on the use of the multi-stage and

two-stage models to describe gastric tumor risk in mice exposed to benzo[*a*]pyrene and lung cancer risk among uranium miners exposed to both radon and tobacco smoke. Both applications involved a wide range of exposure patterns that were taken into account in model fitting. In the former case, the available data suggested that earlier exposures are more important than later exposures with respect to gastric tumor risk. In the latter case, risk assessment methods for time-dependent exposures were extended to the case of exposure to two agents (radon and tobacco), allowing for an assessment of the degree of interaction between these two agents. The age-specific relative risk of lung cancer among miners exposed to both radon and tobacco appeared to be supra-additive but submultiplicative.

Designing informative studies with time-varying doses will be more challenging than designing studies with constant dosing patterns. Although the experimenter has considerable latitude in the choice of exposure regimens, it is not clear how such regimens should be structured in order to provide maximal information on the relationship between time of exposure and risk. One possible objective in developing informative experimental designs with time-dependent exposure patterns would be to obtain accurate estimates of the relative effectiveness function $R(t)$ at different times t.

Studies can also be designed to evaluate the relative effect of different time-dependent exposure patterns. The material herein suggests that, if the model (multi-stage or TSM) is valid, then one should expect the same risk levels for subjects exposed for short, well-defined time periods as those exposed chronically to the appropriately calculated LECD. This presents a quantifiable hypothesis which can be used to explore the validity of the models using experiments with different temporal patterns of exposure.

Although the risk assessment methods described in this article are model based, nonparametric approaches to risk assessment with time-dependent exposure patterns may also be possible. This would permit the use of model-free relative effectiveness functions that may enjoy a degree of robustness against model misspecification. Research in this area is underway.

ACKNOWLEDGMENTS

We wish to thank the reviewers of this manuscript for a number of constructive comments which helped improve the presentation of this material.

REFERENCES

Armitage, P. and R. Doll. 1957. A two-stage theory of carcinogenesis in relation to the age distribution of human cancer. *British Journal of Cancer*. 11:161–169.

Armitage, P. and R. Doll. 1961. Stochastic models for carcinogenesis. *In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 4:19–38, Berkeley, CA. University of California Press.

Chen, J., R. Kodell, and D. Gaylor. 1988. Using the biological two-stage model to assess risk from short term exposures. *Risk Analysis*. 8:223–230.

Crump, K. and R. Howe. 1994. The multi-stage model with a time-dependent dose pattern: application to carcinogenic risk assessment. *Risk Analysis*. 4:163–176.

Day, N. and C. Brown, 1980. Multi-stage models and the primary prevention of cancer. *Journal of the National Cancer Institute*, 64: 977–989.

Denes, J. and D. Krewski, 1994. An exact representation of the generating function for the Moolgavkar-Venzon-Knudson two-stage model of carcinogenesis with stochastic stem cell growth. Submitted to *Mathematical Biosciences*.

Goddard, M. and D. Krewski, 1995 *(In Press)*. The future of mechanistic research in risk assessment: Where are we going and can we get there from here? *Toxicology*.

Kodell, R., D. Gaylor, and J. Chen, 1987. Using average lifetime dose rate for intermittent exposures to carcinogens. *Risk Analysis*. 7:339–345.

Krewski, D., M. Goddard, and J. Zielinski, 1992. Dose-response relationships in carcinogenesis. In H. Vainio, P. Magee, D. McGregor, and A. McMichael, editors, *Mechanisms of Carcinogenesis in Risk Identification*, Lyon. International Agency for Research on Cancer.

Krewski, D. and D. Murdoch, 1990. Cancer modeling with intermittent exposures. In S. Moolgavkar, editor, *Scientific Issues in Quantitative Cancer Risk Assessment*, 196–214, Boston, Mass. Birkhauser.

Moolgavkar, S., 1986. Carcinogenesis modeling: From molecular biology to epidemiology. *Annual Review of Public Health*. 7:151–169.

Moolgavkar, S., G. Luebeck, D. Krewski, and J. Zielinski, 1993. Radon, cigarette smoke, and lung cancer: A re-analysis of the Colorado plateau uranium miners' data. *Epidemiology*. 4:204–194.

Moolgavkar, S. and G. Luebeck, 1990. Two-event model for carcinogenesis: Biological, mathematical, and statistical considerations. *Risk Analysis*. 10:323–341.

Morrison, P., 1987. Effects of time-variant exposure on toxic substance response. *Environmental Health Perspectives*. 76:133–140.

Murdoch, D. and D. Krewski, 1988. Carcinogenic risk assessment with time-dependent exposure patterns. *Risk Analysis*. 8:521–530.

Murdoch, D., D. Krewski, and J. Wargo, 1992. Cancer Risk assessment with intermittent exposure. *Risk Analysis*. 12:569–577.

National Research Council. 1988. *Complex Mixtures, Methods for In Vivo Toxicity Testing*. National Academy Press, Washington, DC.

National Research Council, 1992. *Spacecraft Maximum Allowable Concentrations.* National Academy Press, Washington, DC.

National Research Council. 1993. *Pesticides in children's diets*. In preparation.

Neal, J. and R. Rigdon. 1967. Gastric tumors in mice fed benzo[a]pyrene: a quantitative study. *Texas Reports in Biology and Medicine,* 24:553–557.

U.S. Department of Agriculture. 1978. *National Food Consumption Survey 1977 78.* U.S. Department of Agriculture, Washington, D.C.

U.S. Environmental Protection Agency. 1986. Guidelines for carcinogen risk assessment. *Federal Register*, 51:33992 34003.

Whittemore, A. and J. Keller. 1978. Quantitative theories of carcinogenesis. *SIAM Review,* 20:(1):1–30.

TEMPORAL ASPECTS OF RISK CHARACTERIZATION OF LEAD

Kathryn R. Mahaffey, Ph.D.

Risk Characterization Team, Environmental Criteria Assessment Office
U.S. EPA, 26 West Martin Luther King Drive, Cincinnati, Ohio 45268

INTRODUCTION

Proper selection of risk assessment methodology for lead depends upon the context of the need. For example, at one extreme lies very specifically targeted assessments which consider only one source of exposure (e.g., selecting which surfaces coated with lead-based paint to abate). At the other end lies holistic risk assessments that include consideration of multimedia exposure, as well as evaluation of several sensitive subpopulations and dose-response curves. The latter type of risk assessment is frequently encountered in support of regulations issued by federal environmental agencies. Increased emphasis has been placed on ecological risk in recent decades, although some earlier assessments did include ecological effects of lead (e.g., the U.S. EPA Airborne Lead Criteria Document, 1986). Despite the varied purposes of risk assessments, risk characterization is conducted either formally or informally in the assessment process.

RISK CHARACTERIZATION

Risk characterization is one of the four steps of risk assessment included in the paradigm described by the National Research Council (NRC) of the National Academy of Sciences (1983). Risk characterization is the last step of the four described by NRC (1983): preceded by hazard identification, exposure assessment, and evaluation of dose-response. In response to the requirements of Clean Air Act Amendments of 1990, the NRC/NAS was engaged by US EPA to review risk assessment methods used by US EPA, particularly for chemical contaminants present in air. The report of the NRC committee that revisited the risk assessment process, *entitled Science and Judgment in Risk Assessment* (NRC/NAS, 1994), indicated that the risk characterization process consisted of four elements: generation of a quantitative estimate of risk, qualitative description of uncertainty, presentation of the risk estimate, and communication of the results of risk analysis. A risk characterization should delineate the balance between confidence and uncertainty in various steps in the assessment process. A qualitative description of the elements that form the basis of numerical estimates

produced in the risk assessment must always accompany numerical estimates if users of the risk characterization are to fully grasp the strengths and weaknesses. The risk characterization process must also consider alternatives to the one chosen. The strengths and weaknesses of each alternative approach should be described. Variability around estimates of central tendency (e.g., percentiles, range, variance) are essential to risk characterization in order to permit decision makers to estimate the strength of the association and the likely impact of risk reduction measures.

Risk assessment can be done for an agent (physical, chemical or biological condition), as well as for a geographic location (e.g., a Superfund site or a residential unit). Agent-specific risk assessments preferably evaluate all routes of exposure: inhalation, ingestion, and dermal. These routes or pathways of exposure reflect intake from multiple media: air, food, water, dust, soil, etc. These routes of exposure result from primary sources that may be of natural or anthropogenic origin. The latter source category includes: mining, smelting, emissions from industrial activities, leaching of man-made chemicals into water, etc. Risk assessment for a geographic location or site-specific characterizations focus on multiple agents at a particular location (as with a Super-fund site) or multiple locations for a single chemical (as with residential assessment of lead-based paint hazards in federally assisted housing).

TEMPORAL ASPECTS OF RISK CHARACTERIZATION FOR LEAD

This discussion describes changes in the level of certainty over the question of subtle neurotoxicity following low levels of lead exposure. It also considers how temporal patterns in lead exposure contribute to variation in expression of these lead effects. In this analysis of temporal characteristics, it is important to separate information on "what was known at the time of the assessment" from time-related differences in lead exposures and the production of adverse health effects. The temporal component "what we knew at the time of the assessment" really describes the degree of certainty at a particular point in time. The rapid proliferation of data on health consequences of lead exposures coupled with an extraordinary decrease in lead exposure during the past two decades has produced major shifts in our certainty/uncertainty of the health consequences and the significance of sources/pathways of lead exposure. Major questions that are part of current risk characterization are shown in Table V-10.

Table V-10. Temporal Components in Risk Characterization for Lead.

Element of Risk Assessment	Component of Risk Characterization		
	Assumption	Uncertainty	Variability
Hazard Identification	Neurotoxicity most sensitive health-based end-point. Fetus and young child most susceptible to neurotoxicity.	Is the effect reversible? Is there a threshold for this effect? What is the critical exposure period?	Are some members of the population at much greater susceptibility than others for biological reasons?
Exposure Assessment	Lead is a cumulative toxin. Time and intensity of exposure can be modeled or measured using currently available methodology.	How much past exposures (i.e., non-contemporaneous) contribute to biological indicators of exposure (e.g., blood lead concentrations)? How to retrospectively estimate exposures?	How much do chronological changes in exposure patterns modify exposure estimates?
Dose–Response Evaluation	That the lowest end of the dose–response curve has been identified.	If dose were reduced would the range of effects be expanded? That is, if exposures already occur at levels high enough to produce effects would conditions perceived as "normal", in fact, be effects?	What differences exist among dose–response curves for specific subpopulations?

RISK CHARACTERIZATION FOR PEDIATRIC LEAD TOXICITY

The history of pediatric lead poisoning in the United States is marked by a shift from case finding to a focus on primary prevention of lead exposure. In the first half of the 20th century, most public efforts were focused on case reporting and creation of registries of cases. The first mass screenings of pediatric populations to identify high-risk children began in Chicago in 1966 followed shortly by screening in New York City, Baltimore and Philadelphia (Lin-Fu, 1970; Wolman, 1971). The seriousness of the problem was illustrated by the New York City data. During the period January–October (1970) about 55,000 children were tested; 3% of these children had blood lead concentrations

above 60 mg/dL and 45% above 40 mg/dL (Wolman, 1971). In 1971, the Surgeon General issued a statement emphasizing the need to shift the focus from identifying poisoned children to primary prevention and the United States Congress passed the Lead-Based Paint Poisoning Prevention Act (Public Law 91-695, 91st Congress, HR 19172, 13 January 1971) emphasizing primary prevention of lead exposures from housing. The use of blood lead concentration to express exposure, as well as a metric of health effects of lead has been significant to subsequent risk assessments: virtually every risk assessment during the 1960s and later has been structured around blood lead concentrations (NRC, 1993).

NEUROTOXICITY FOLLOWING CHILDHOOD EXPOSURE TO LOW LEVELS OF LEAD: INCREASED CERTAINTY OF THE EFFECT

During the first half of this century in the United States, clinical morbidity (such as, profound mental retardation and seizures) and death were how lead hazards were identified (Williams, 1952; Lin-Fu, 1970; Lin-Fu, 1985; Perlstein and Attala, 1966). Only with increased understanding of how organ systems were affected by lead were subtle effects of lead on health identified. This increased understanding was accompanied by a shift from secondary to primary prevention. A shift from case finding/registries to prevention was illustrated by publication of the report by National Research Council of the National Academy of Sciences, *Lead: Airborne Lead in Perspective* in 1972. This NRC Committee's summary of data on overt cases of lead poisoning included severe lead effects on the nervous system; such as, encephalitis and marked intellectual retardation among survivors of severe lead poisoning. However, this NRC Committee's report (which can be regarded as an early risk assessment) indicated that more subtle effects of lead were identified through the hematological system. Whether or not behavioral or learning deficits could be attributed to lead exposure was considered by this NRC Committee, but in 1972 whether or not neurobehavioral or neurotoxicity occurred at lead exposures lower than those producing hematological effects was a completely unresolved issue. In the early 1970s, impairment of hematopoiesis, as shown by biochemical indicators of heme biosynthesis, was regarded as the most subtle indicator of lead hazard. The predominance of hematological end-points in the early 1970s also can be recognized in the Centers for Disease Control's guidelines (1975, 1978) that were based on erythrocyte protoporphyrin concentrations and blood lead levels.

Although the hematopoietic system was regarded as the organ system most sensitive to lead effects, research on the neurological effects of lead continued. A series of cross-sectional studies,

reported during the 1970s, suggested that subtle learning deficits were associated with blood lead concentrations in the range of 20 to 40 mg/dL. The problems inherent to cross-sectional studies were addressed in the longitudinal prospective studies that included: the Cincinnati Longitudinal study (Dietrich et al.; 1987, 1991, and 1993), the Port Pirie study (Baghurst et al., 1992), the Boston study (Bellinger et al., 1987; 1988), the Cleveland study (Ernhart et al., 1987; 1989), the Sydney study (Cooney et al., 1989), the Kosovo study (located in the former Yugoslavia) (Wasserman et al., 1992). In the longitudinal studies, typically, recruitment began with pregnant women; the infant was evaluated repeatedly between birth and, at least, early school age. The cohorts differed substantially in their socio-economic status, and the presence of important covariables, such as alcoholism in a substantial portion of the mothers from Cleveland, low-iron status in the Kosovo population, or rearing environment in the Cincinnati cohort.

There is a negative association between increasing levels of lead exposure (expressed as increases in blood lead concentrations) and decreasing scores for global indices of intellectual development or neuropsychological function, typically represented as IQ. Most analyses of the data included assessment of global IQ as the index of the neurotoxicity of lead. However, the neurotoxic effect is complex (see Bellinger, in press) owing to either intrinsic individual differences, age at which lead exposures occurred or numerous confounding variables such as nutritional status, concurrent disease, and exposures to other toxic chemicals. In some cohorts the neuropsychologic effects most associated with lead are in the verbal domain, while in other cohorts the effects are most strongly associated with the nonverbal domain.

Over the next 15-year period (approximately 1980 through 1994) the association between subtle delays and deficits in intellectual development and lead exposure were further clarified. The presence of a lead effect and the magnitude of the effect differed among the cohorts. Using meta-analyses based on cross-sectional data, Needleman and Gatonis (1990) estimated the 95% confidence interval of the correlation between blood lead and IQ to be –0.1 to –0.2. Following analysis of the combined cross-sectional and prospective studies, Schwartz (1994) estimated the decrease in IQ score for a 10 mg/dL increase in blood lead to be approximately 2.6 points (95% CI: 1.8 to 3.4). Pocock et al. (1994) using meta-analysis techniques estimated that an increase in blood lead concentration from 10 to 20 mg/dL would be associated with a mean deficit in full scale IQ of around 1 to 2 IQ points. Several "official" risk assessments were prepared by the United States government (e.g., CDC, 1991; EPA 1989). The major change in these assessments was increased certainty that the neurobehavioral deficits were

associated with lead exposures common in the general population of infants and young children in the United States. In response to this increased certainty of neurobehavioral effects of low level lead exposures, the Centers for Disease Control's definitions of lead poisoning were revised in 1985 and again in 1991.

The strength of the associations between blood lead and results of the neurobehavioral tests changed as the age of the children in the longitudinal cohorts increased. Data on the strength of the association between increases in blood lead concentration and decrements in intelligence based on testing during early childhood (e.g., 2 and 3 year-olds) varied between cohorts. Identification of an association was not consistent (contrast Bellinger et al., with Ernhart et al.) and among those studies showing an association the magnitude differed (Bellinger et al., Dietrich et al., etc.). However, those longitudinal studies that continued testing into early grammar school more consistently identified impaired performance on tests of intelligence. Lead exposures resulting in blood lead concentrations over the range of 10 to 30 mg/dL are associated with an intellectual decrement of approximately 2 to 4 IQ points per 10 mg/dL increase in blood lead.

Currently three of the cohorts continue to be followed: Cincinnati, Boston, and Port Pirie. It is anticipated that follow-up assessments will be carried out over the next decade. Such research will assist in determining whether or not the early deficits associated with lead exposure are either "reversible" physiologically or compensatable because of plasticity of the nervous system, or result in permanent deficits with attendant diminution in life-long success. Nevertheless, the presence of learning deficits during developmentally critical periods is an adverse effect.

VARIABILITY AMONG SUBPOPULATIONS OF CHILDREN EXPOSED TO LEAD

Temporal characteristics of exposure assessment include: years in which the exposure was estimated, timing of exposure with respect to adverse health effects, timing of sampling of environmental media, and time-course as it modifies dose-response to lead. Quantitatively associating environmental lead exposure with adverse health effects in children is complex. The temporal changes in blood lead concentrations with time of exposure, age of the subject, or effects of other conditions (e.g., marginal nutritional status, infectious disease) must be considered to understand dose-response to lead. When neurotoxicity is the basis for health hazard, the developmental period in which susceptibility to the adverse effects of lead exposure is highest becomes a critical consideration in assessing dose-response to lead. However, the longitudinal "tracking" of blood lead levels precludes a clear-cut identification of such a critical period of developmental susceptibility.

Dose for lead can be considered as either environmental (external) or internal (tissue concentrations) dose. In various risk assessments conducted during the past decades, the majority expressed internal dose as blood lead concentration. Blood lead concentration has been assumed to reflect lead transferred from the concurrent or recent environment (i.e., external dose). Although this has been the traditional interpretation, current studies among women and children have begun to identify the magnitude of the contribution of body stores of lead to blood lead concentrations. For example, Gulson et al. (*in press*) have determined that among adult women of child-bearing age (having blood lead concentrations under 20 mg/dL) approximately 45 to 75% of the blood lead is derived from body stores of lead; that is, lead from previous rather than contemporaneous environments. The magnitude of this effect for young children is not yet known.

Blood lead concentrations have been used as a biomarker of exposure to lead. For groups of individuals, blood lead concentrations reflect environmental exposure over a moderate period of time. During the past 15 years there has been a major decline in environmental lead exposure in the United States. For example, the second National Health and Nutrition Examination Survey (NHANES II), conducted between 1976 and 1980, reported that the mean blood lead concentration for all races and all ages in the United States population was 13.9 mg/dL (Mahaffey et al., 1982). The distribution of blood lead concentrations was highly uneven: 22% of the population had blood lead concentration of < 10 mg/dL, whereas 1.9% across all age groups had levels of^3 30 mg/dL. Phase I of the third National Health and Nutrition Examination Survey (NHANES III) conducted between 1988 and 1991, identified the mean blood lead concentration for the United States population as 2.8 mg/dL (Brody et al., 1994). This decrease of approximately 80% from the levels reported 12 years earlier (Mahaffey et al., 1982) reflected a concerted public health and environmental health effort to reduce use of lead in the United States. Reductions in environmental lead reflect the virtual elimination of the use of lead-solder in food and beverage cans, the removal of lead additives from gasoline, and strong public health programs independent of source reduction.

However, the uneven distribution of elevated blood lead levels continues. Children are the subpopulation considered at greatest risk of adverse effects of lead. In the NHANES II data set, 18.5% of African-American children from low-income, inner-city families had blood lead levels ^{3} 30 mg/dL (Mahaffey et al., 1982). The NHANES III data showed that 4.5% of all one-to-two year-old children had blood lead levels over 10 mg/dL; however 21.6% of one-to-two year-old children of African-American ancestry had blood lead concentrations over 10 mg/dL (Brody et al., 1994).

The substantial change in lead exposure for the general population modifies risk characterizations for lead. Understanding the range of adverse health effects produced by lead requires the risk assessor to recognize that the elevated blood lead concentrations observed in the 1970s were sufficiently high that subtle effects of lead on neurobehavioral development were difficult to separate from normal "background" neurobehavioral deficits. For example, 63.3 percent of low-income children (ages 6-months-through-5-years) of African-American descent living in inner-cities in the years 1976 to 1980 had blood lead concentrations > 20 mg/dL (Annest and Mahaffey, 1984). In 1990, the Centers for Disease Control recommendations were that children with blood lead concentrations of 20 mg/dL be referred for medical management for lead poisoning (CDC, 1991). Only with decreased environmental exposures have the adverse health effects of continued high exposure become more certain.

TEMPORAL COMPONENTS IN DOSE-RESPONSE TO LEAD: TIMING OF EXPOSURE IN RELATION TO HEALTH EFFECTS

Choice of the health end-point to use in the risk assessment will determine what exposure periods are particularly important. For example, when hematological end-points were used as critical health effects (the situation in virtually every risk assessment prior to 1980), recent lead exposures (those within the past two to three months) were thought to be the ones to consider because of continuous synthesis of new heme and the half-life of the erythrocyte. By contrast, the critical period for lead-related effects on CNS development are thought to be late gestation through early childhood (Rodier, 1986; Verity, 1990; Goldstein, 1990). Consequently, knowledge of lead exposures during this developmental period is pivotal to grasping the association between environmental/occupational lead exposure and CNS-based health outcomes. To the extent that lead is recycled internally as part of the biokinetics of normal human mineral metabolism, long-term exposures from gestation onward contribute to development of adverse neurobehavioral effects of lead.

Another temporal complication in developing dose-response relationships for children is the episodic nature of the highest lead exposures that occur secondary to pica. Pica refers to ingestion of non-food substances that may include lead-containing paints or soils. These ingestion patterns typically are not regularly timed, in contrast to lead exposures from media such as air or food. Experience with efforts to model the dose-response aspects of this pica-dependent exposure identified kinetics that differ from those following more routinely ingested lead sources (LaKind et al., 1994).

TEMPORAL COMPONENTS OF DOSE-RESPONSE TO LEAD: TIMING OF SAMPLING OF ENVIRONMENTAL MEDIA

Linking estimates of environmental lead exposures to development of adverse health effects of lead is complex because lead toxicity can be acute (in onset) or a chronic condition reflecting years of accumulated excess lead exposures. Knowledge of temporal aspects of lead toxicity (e.g., determining the time period when exposures are of greatest biological impact) helps to target environmental sampling for lead. Presence of an internal "source", tissue lead burden, that contributes to blood lead increases the importance of tissue accumulation of lead which continues to be recycled to organs such as brain.

Short-term high exposures such as lead ingestion of lead-based paint chips by an infant or from house renovations involving lead-based paints are easier to associate with increases in blood lead. Choice of appropriate sampling strategies in these examples are more clear-cut; ideally based on well-developed principles of industrial hygiene or environmental exposure assessments. Sampling environmental media in situations where lead toxicity is due to steady accumulation of small excesses in lead exposure over time is a different situation. Most data on this process have been obtained in the longitudinal prospective studies associating lead exposures with intellectual deficits among children. One observation from these studies is that individual children of a cohort after approximately two-years of age begin to maintain their rank-order in the cohort; i.e., the child's relative position in the cohort if blood lead concentrations are ranked from lowest to highest.. This process may reflect an individual child's interaction with his/her environment, remobilization of earlier body burden of lead, genetically based individual differences in distribution of a given body burden of lead across different tissue pools, or some combination of these. Sampling method and sampling location are also critical in conducting residential lead risk assessments (Lamphear, *in press*).

Any effort to quantitatively estimate lead exposure needs to address the concern that the lead measured in environmental samples be representative of the lead ingested/inhaled by the subject. Sampling strategies need to address spatial, as well as temporal, components. To the extent that the sampling methodology does not adequately identify lead ingested by a young child during a critical period, the strength of the association between environmental/occupational lead exposures and either health end-points or blood lead concentration is weakened. The so-called "errors in the variables" problem is described in numerous statistical texts (among other see, Draper and Smith, 1981). To the

extent that these errors occur, the association between environmental lead exposure and health-based end-points will be underestimated.

CONCLUSIONS

Risk characterization for the effect of low level lead exposure on neurotoxicity among young children recognizes greatly increased certainty that subtle intellectual deficits are associated with lead exposure among children. Much of the current dialogue in risk characterization focuses on the magnitude of this effect in the general population; i.e., variability. Efforts to prevent childhood lead exposure result in reducing the quantities of lead that a child ingests/inhales from environmental sources. The pathways of lead exposure are qualitatively known. "State of the art" assessment must focus on the relationship between blood lead and spatial and temporal patterns in environmental media.

REFERENCES

Annest, J.L. and Mahaffey, K.R. 1984. Blood Lead Levels for Persons Ages 6 Months – 74 Years: United States, 1976–1980. *Vital and Health Statistics*. U.S. Department of Health and Human Services, Public Health Service. National Center for Health Statistics. Series 11, No. 233. Hyattsville, Maryland.

Baghurst, P., McMichael, A., Wigg, N., Vimpani, G., Robertson, E., Roberts, R., and Tong, S-L. 1992. Environmental exposure to lead and children's intelligence at the age of seven years. *N.Engl. J. Med.* 327:1279–1284.

Bellinger, D. (1995, in press) Interpreting the literature on lead and child development: The neglected role of the "experimental system". Open Peer Commentary Perspective. Neurotox. & Teratol.

Bellinger, D., Leviton, A., Waternaux, C., Needleman, H., and Rabinowitz, M. 1987. Longitudinal analyses of pre- and postnatal lead exposure and early cognitive development. *N. Engl. J. Med.* 316:1037–1043.

Bellinger, D., Leviton, A., Waternaux, C., Needleman, H., and Rabinowitz, M. 1988. Low-level lead exposure, social class, and infant development. Neurotox. & Teratol. 10:497–503.

Brody, D.B., Pirkle, J.L., Kramer, R.A., Flegal, K.M., Matte, T.D., Gunter, E.W., and Paschal, D.C. 1994. Blood lead levels in the U.S. population. Phase 1 of the Third National Health and Nutrition Examination Survey (NHANES III, 1988 to 1991). *J. Amer. Med. Assn.* 272:277–283.

CDC (Centers for Disease Control). 1975. Increased Lead Absorption and Lead Poisoning in Young Children. Atlanta, GA: CDC, U.S. Department of Health, Education, and Welfare.

CDC (Centers for Disease Control). 1978. Preventing Lead Poisoning in Young Children: A Statement by the Centers for Disease Control. DHEW Publ. No. 00-2629. Atlanta, GA: CDC, U.S. Department of Health and Human Services.

CDC (Centers for Disease Control). 1985. Preventing Lead Poisoning in Young Children: A Statement by the Centers for Disease Control. January, 1985. DHHS Publ. No. 99-2230. Atlanta, GA: CDC, U.S. Department of Health and Human Services.

CDC (Centers for Disease Control). 1991. Preventing Lead Poisoning in Young Children: A Statement by the Centers for disease control — October 1991. Atlanta, GA: CDC, U.S. Department of Health and Human Services.

Cooney, G., Bell, A., McBride, W., and Carter, C. 1989. Low-level exposures to lead: the Sydney Lead Study. *Dev. Med. Child Neuro.* 31:640–649.

Dietrich, K., Krafft, K., Bornschein, R., Hammond, P., Berger, O., Succop, P., and Bier, M. 1987. Low-level fetal lead exposure effect on neurobehavioral development in early infancy. *Pediatr.* 80:721–730.

Dietrich, K., Succop, P., Berger, O., Hammond, P., and Bornschein, R. 1991. Lead exposure and the cognitive development of urban preschool children. The Cincinnati Lead Study cohort at age 4 years. *Neurotox. & Teratol.* 13:203–211.

Dietrich, K., Berger, O., Succop, P., and Hammond, P. 1993. The developmental consequences of low to moderate prenatal and postnatal lead exposure: Intellectual attainment in the Cincinnati Lead Study cohort following school entry. *Neurotox. & Teratol.* 15:37–44.

Draper, N.R., and Smith, H. 1981. *Applied Regression Analysis,* Second Edition. John Wiley & Sons Publishers. New York.

Ernhart, C., Morrow-Tlucak, M., Marler, M., and Wolf, A. 1987. Low level lead exposure in the prenatal and early preschool periods: Early preschool development. *Neurotox. & Teratol.* 9:259–270.

Ernhart, C., Morrow-Tlucak, M., Wolf, A., Super, D., and Drotar, D. 1989. Low level lead exposure in the prenatal and early postnatal periods: Intelligence prior to school entry. *Neurotox. & Teratol.* 11:161–170.

Goldstein, G.W. 1990. Lead poisoning and brain cell function. *Environ. Health Perspect.* 89:91–94.

Gulson, B., Mahaffey, K., Cameron, M., and Mizon, K. (in press, 1995) Contribution of tissue lead to blood lead in adult female subjects based on stable lead isotope methods. *Journal of Laboratory and Clinical Medicine, in press.*

LaKind, J.S., Youngern, S.H., Piccin, T.B., and Naimar, D.Q. 1994. Comparison of four models for predicting blood lead levels in children. Abstract No. S21.05. In: Final Program. Society for Risk Analysis. p. 40. December 4–7, 1994. Annual Conference of the Society for Risk Analysis. Baltimore.

Lin-Fu, J.S. 1970. Childhood lead poisoning: an eradicable disease. *Children* 17:2–9.

Lin-Fu, J.S. 1985. Historical perspective on health effects of lead. pgs. 43–63. In: *Dietary and Environmental Lead: Human Health Effects.* Ed. Mahaffey KR. Elsevier Science Publishers, B.V. (Biomedical Division). Elsevier Publishers, Amsterdam/New York.

Mahaffey, K.R., Annest, J.L., Roberts, J., and Murphy, R.S. 1982. National estimates of blood lead levels: United States, 1976–1980. *N. Engl. J. Med.* 307:573–579.

National Research Council. 1972. *Airborne Lead in Perspective.* National Academy Press. Washington, DC.

National Research Council. 1983. *Risk Assessment in the Federal Government.* National Academy Press. Washington, DC.

National Research Council. 1993. *Measuring Lead Exposure in Infants, Children and other Sensitive Populations.* National Academy Press. Washington, DC.

National Research Council. 1994. *Science and Judgment.* National Academy Press. Washington, DC.

Needleman, H., and Gatsonis, C. 1990. Low-level lead exposure and the IQ of children. *J. Amer. Med. Assn.* 263:569–573.

Perlstein, M.A., and Attala, R. 1966. Neurologic sequelae of plumbism in children. *Clin. Pediatr.* 5:292–298.

Pocock, S., Smith, M., and Baghurst, P. 1994. Environmental lead and children's intelligence: a systematic review of the epidemiological evidence. *Br. Med. J.* 309:1189–1197.

Rodier, P. 1986. Time of exposure and time of testing in developmental neurotoxicology. *NeuroToxicology* 7:69–76.

Schwartz, J. 1994. *Environ. Res. – in press.*

United States Environmental Protection Agency. 1986. Air Quality Criteria for Lead. Environmental Criteria and Assessment Office. Research Triangle Park, NC. EPA-600/8-83/028aF.

Verity, M.A. 1990. Comparative observations on inorganic and organic lead neurotoxicity. *Environ. Health Perspect.* 89:43–48.

Wasserman, G., Graziano, J., Factor-Litvak, P., Popovac, D., Morina, N., Musabegovic, A., Vrenezi, N., Capuni-Paracka, S., Lekic, V., Preteni-Rejepi, E., Hadzialjevic, S., Slavkovich, V., Kline, J., Shrout, P., and Stein, Z. 1992. Independent effects of lead exposure and iron deficiency anemia on developmental outcome at age 2 years. *J. Pediatr.* 121:695–703.

Williams, H., Kaplan, E., Couchman, C.E., and Sayres, R.R. 1952. Lead poisoning in young children. *Publ. Hlth. Reports* 67: 230–236.

Wolman, I.J. 1971. The Federal attack on childhood lead poisoning. *Clin. Pediatr.* 10: 692–693.